The University Of Sheffield.

A Study of Process Utility.

Volume 1

By:

Victoria K Brennan

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Medicine, Dentistry and Health
School of Health and Related Research

October 2015

A Study of Process Utility.

Victoria K Brennan

Abstract

Background: This study compares the cost-effectiveness of using an online assessment tool (ePAQ) in advance of a face-to-face (control) versus a telephone (intervention) consultation. The trial suggests statistically significant differences between groups for three PEQ (Patient Experience Questionnaire) dimensions. However, it is unlikely that these differences were captured by the cost-effectiveness outcomes (Incremental cost/QALY). This is consistent with other work that has suggested the QALY does not capture all outcomes relevant to interventions, particularly where the outcomes are non-health outcomes e.g. associated with processes of receiving care.

Aim: With this issue in mind, the ePAQ trial was a vehicle for the exploration of process utility. The overarching aim was to utilise ePAQ trial data, to improve the relevance of the cost-effectiveness outcomes by incorporating process utility into the QALY.

Methods: This thesis includes: a literature review of process utility studies; a process utility bolt-on study; the application of the findings to ePAQ trial data; revised cost-effectiveness estimates incorporating process utility. Qualitative methods were used to explore participant's reactions to health states including health and process outcomes.

Results: The valuation study captured process utility. The inclusion of process utility in the ePAQ EEACT moved the intervention from being dominated by the control, to a position where it falls within the willingness-to-pay threshold of £20,000/QALY, indicating its inclusion can impact on cost-effectiveness results. Respondents' were able to value health and process scenarios within the think aloud study.

Conclusion: This study indicates that patients do care about both health and process outcomes. The valuation study provided a mechanism for capturing the amount of utility people have for these processes, however, the qualitative study raised questions on the validity of utility estimates derived from health and process scenarios. This thesis is an exploratory study, and highlights a need for further research into the topic.

Contents

List of Figures.

List of Tables

Acknowledgements.

Thanks to my family, especially my parents.

To my Mum who has continually told me I am as good as everyone else, and to have confidence and belief in myself. And to my Dad, for the values he has instilled in me and for taking care of my aging horse in my absence made more by work. Thank you both for providing me with an education I feel privileged to have had, and the opportunities that has brought. And thank you for instilling in me the importance of independence, education and the desire to learn and succeed.

Thank you to my siblings, for always being there. To my Grandpa for inspiring me and to my Nanna for always being a voice of calm. I miss you.

Thank you to Tony, for always believing in me and being my calm in my many moments of worry and panic, and for never letting me give up.

A huge thank you to my supervisors, for helping me to grow in knowledge and confidence, and for supporting me through the ups and downs of the last 8 years of study.

Thank you to Mr Fisher, my A level Economics teacher, who introduced me to a subject that sparked a life-long interest. To RTI for giving me my first job as a health economist and for funding my studies, and to Cook Medical, for allowing me the time to write up my thesis.

I would also like to thank Brendan Mulhern for your ongoing support and advice. Rob White and the team from Epigenesys for all of your work in developing the online survey, and SSI for your work in hosting the survey. Thanks to the ScHARR data collection fund for funding the survey. Thanks also to Jo Thomas and Abu Alshreef for your time and expertise.

And finally to Sam, for always being there, and for giving me a moment of calm and tranquillity whenever I need it.

Chapter 1: Thesis Overview

1.1 Background

This thesis focuses on the traditional methods of economic evaluation used in the UK. These are cost-utility analyses (CUA) which use the quality adjusted life year (QALY) as a measure of benefit (Brazier et al. 2007). This approach is now required for cost-effectiveness analyses by the National Institute for Health and Care Excellence (NICE), and the All Wales Medicines Strategy Group (AWMSG), as well as other national decision making bodies such as the Canadian Agency for Drugs and Technologies in Health (CADTH), and the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia.

Within this thesis the standard economic framework for undertaking CUA is applied in an economic evaluation alongside a clinical trial (EEACT) undertaken in pelvic floor medicine (ePAQ EEACT). The ePAQ EEACT compares the use of an online assessment tool (ePAQ), in advance of either a face-to-face (control) or telephone (intervention) consultation, in women with pelvic floor problems. The primary economic outcome is an incremental cost per QALY.

This thesis is based on one fundamental issue with the ePAQ trial study design. The trial essentially compares the consultation approach (face-to-face versus telephone) between groups. This, we can assume, will have no impact on health outcomes, but

that any differences between groups will be due to patients' satisfaction for the different processes of receiving care. This is confirmed by the choice of the primary clinical outcome measure, the Patient Experience Questionnaire (PEQ) which is a measure of patient satisfaction. Within traditional economic evaluation, the QALY is the preferred health economic outcome (Drummond, 2005). However, the most commonly used and favoured preference based measures which are used to derive the utilities for input into the QALY calculation, tend to focus purely on health effects, and fail to consider other non-health attributes such as those associated with the "process" of receiving care (Donaldson and Shackley, 1997). The issue with the ePAQ study design therefore, is that the QALY in its current form is only designed to capture differences in preferences for health outcomes. It is not designed to capture patients' preferences for the processes of receiving care – as it does not incorporate process utility (Mooney, 1998).

With this fundamental issue in mind, the ePAQ EEACT was used as a vehicle for the exploration of the presence and measurement of process utility. The overarching aim of the thesis was to utilise data collected within the ePAQ trial, to improve the relevance of the cost-effectiveness outcomes through the incorporation of process utility into the QALY calculation.

1.2 Aims and Objectives

1.2.1 Aims

This research aims to use the ePAQ EEACT to explore process utility in terms of its presence, its measurement, and its incorporation into the QALY calculation. The over-riding aim of the thesis is to improve the relevance of the cost-effectiveness outcomes from the ePAQ EEACT.

1.2.2 Objectives

The objectives of this study are:

- To perform the cost-effectiveness analysis for the ePAQ EEACT;
- To identify current views on process utility, and the methods which have been used in previous studies to identify and measure process utility;
- To design and carry out a study for the valuation of process utility, which can be incorporated into the QALY calculation.
- To apply the findings from the process utility valuation study to the ePAQ EEACT, and produce alternative cost-effectiveness estimates incorporating process utility.

- To assess the validity of process utility values captured by the time trade-off technique.

1.3 Structure of the Thesis

Chapter Two provides an introduction to the traditional methods of economic evaluation. It includes background on the need and types of economic evaluations. It then focuses on cost-utility analyses, specifically those which use the QALY as a measure of benefit. The QALY is introduced, and the outcomes that it encapsulates are discussed. Finally the chapter reports a literature review into the limitations of economic evaluations alongside clinical trials (EEACT). The background information reported in this chapter provides a context for the subsequent research reported in the thesis.

Chapter Three applies the traditional methods of economic evaluation in practice. It reports an EEACT which was undertaken at Sheffield Teaching Hospitals (STH), to determine the cost-effectiveness of using an online questionnaire (ePAQ) in combination with a telephone consultation compared to standard care. It concludes by challenging the assumption that individuals only derive utility from the consequences of health care, and suggests that when conducting EEACTs which include processes of receiving care (e.g. a telephone versus a face-to-face doctor's consultation), as opposed to a specific treatment intervention, the utility gained or lost through these processes, which is termed as "process utility" should also be considered, and where appropriate incorporated into the QALY calculation.

Chapter 4 focuses on process utility. Having established that process utility may be an important consideration in the economic evaluation of interventions that alter the process of care. This chapter reports a systematic literature review of studies reporting empirical measures of process utility that could be incorporated into the QALY calculation. The studies are described, including the methodology used to identify and measure process utility, and the recommendations for future research. These sources are used in the subsequent chapters of this thesis to inform the design of the process utility bolt-on study, which aims to utilise data already collected from within the ePAQ EEACT.

Chapter 5 reports the development of the methodology for the process utility bolt-on study. This begins with a summary of a targeted literature review of bolt-on studies within health economics, and then combines the knowledge gained from these studies with knowledge from the previous process utility review reported in Chapter 4 to describe the development of the process utility bolt-on study design. This includes exploratory statistical and principal component analysis performed on the ePAQ study data to define the process domain, a pilot study to format the domain and the development of an online survey to value the health and process scenarios.

Chapter 6 reports the results and findings from the valuation study, and its application back to the ePAQ EEACT study results. It presents the cost-effectiveness results reported in the Chapter 3 (excluding process) alongside the results once process utility

has been accounted for. An updated incremental cost effectiveness ratio (ICER) and cost-effectiveness acceptability curve (CEAC) is shown.

Chapter 7 draws qualitative work into the thesis, and uses a think aloud technique to explore participant's reactions to the health scenarios that include both health and process outcomes, and their ability to value them using recognised and preferred techniques (time-trade off). It initially reports a review of the think aloud literature in health economics. The think aloud study is then reported.

Chapter 8, the final chapter commences with a summary of the research. It provides an overview of the research contributions made to methods for measuring and incorporating process utility into the QALY. The strengths and limitations of the study are reported, and recommendations for future research are made. The thesis concludes by discussing the implications of the findings from this study on cost-utility analyses which are performed to inform resource allocation and policy decisions in health care.

1.4 References

- Brazier, J. Ratcliffe, J. Salomon, JA. Tsuchiya, A. (2007) *Measuring and valuing health benefits for economic evaluation.* Oxford, Oxford University Press.

- Donaldson, C. & Shackley, P. 1997. Does "process utility" exist? A case study of willingness to pay for laparoscopic cholecystectomy. *Soc Sci Med*, 44, 699-707

- Drummond MF, Schulpher MJ, Torrance GW, O'Brien B, Stoddart, GJ (2005) *Methods for the Economic Evaluation of Health Care Programmes.* Third Edition. Oxford, Oxford University Press.

- Mooney, G. (1998) The benefits of health care. *Health Care Analysis*, 6, (2) 99-101.

Chapter 2: Background to Economic Evaluations.

2.1 Introduction

This chapter provides an introduction to the traditional methods of economic evaluation. It includes background on the need for economic evaluations, different types of economic evaluation, the quality-adjusted life-year (QALY) and the outcomes that it encapsulates. It introduces "utilities", which are required for input into the QALY, and asks whether, in addition to health outcomes, outcomes associated with the process of receiving care or "process utility" should also be considered. Alternative ways of conducting cost-utility analyses are then outlined.

Finally this chapter reports a literature review on the limitations of economic evaluations alongside clinical trials (EEACTs). The information in this chapter provides a context for the subsequent research reported in this thesis.

2.2 Economic Evaluations.

Health care resources are finite, and as populations continue to grow and life expectancy extends, the demand for these resources increases. This scarcity combined with the high, and increasing demand for health care, means that we are not able to provide all of the health care services to all of the people who need or request them. Therefore decisions need to be made on how resources are allocated to different

services and patients, and which services or treatments are sacrificed and have resources directed away from them. With this necessity to maximise health gain within a given budget constraint, health needs to be measured and valued and this is done through economic evaluation.

2.2.1 What is an economic evaluation?

Economic evaluation is the comparative assessment of the costs and benefits of alternative health care interventions (Drummond, 2005). There are four main types of economic evaluation, and these are distinguished by the way the benefits are considered (Brazier et al. 2007). A cost-minimisation analysis compares interventions in which the benefits are identical and therefore only cost differences are considered. Cost-effectiveness analyses consider both the cost of each intervention, and a relevant common clinical outcome. A cost-benefit analysis compares the benefits of interventions, where all are valued in monetary terms. An example of this is "willingness to pay" (WTP) where methods are used to set a monetary value for health and/or non-health benefits associated with an intervention (Cookson, 2003). Finally, a cost-utility analysis compares the benefits of the interventions using a measure of utility. This refers to a single value which reflects society's preference for a health state or condition (Drummond, 2005). Most commonly, the utility value is combined with the time spent in the health state to calculate a quality-adjusted life-year (QALY). The

incremental cost per QALY gained is then calculated (Drummond 2005, Brazier et al. 2007). Within England, NICE make recommendations on methods for technology appraisal (NICE, 2013), which provide specific guidance on their preferred approaches for undertaking economic analyses, this includes for example, that health benefits are expressed in QALYs, derived using utilities obtained from public preferences using a choice-based method, preferably time-trade off (TTO) or standard gamble (SG). Hence, these methods are embedded in the current resource allocation decision making processes, and the dominant methods of economic evaluation used today. They are therefore the focus of this thesis. This focus on standard and recommended methods for determining utility excludes methods such as willingness to pay (WTP), which are generally used to determine a monetary value for the health benefits gained by an intervention (Cookson, 2003). Although it is possible to link WTP to utility (Robinson et al. 2013) it was felt that the focus on standard QALY approaches would result in findings more applicable to real-life resource allocation decisions today.

2.2.2 Cost-utility analyses

Cost-utility analyses (CUA) are central to this thesis, and therefore will be discussed in greater depth. CUA's consider both the quantity and quality of life years gained by interventions. The QALY uses a utility value to weight a patient's quality of life within one year, these can be added up to determine an individual's preference for a health

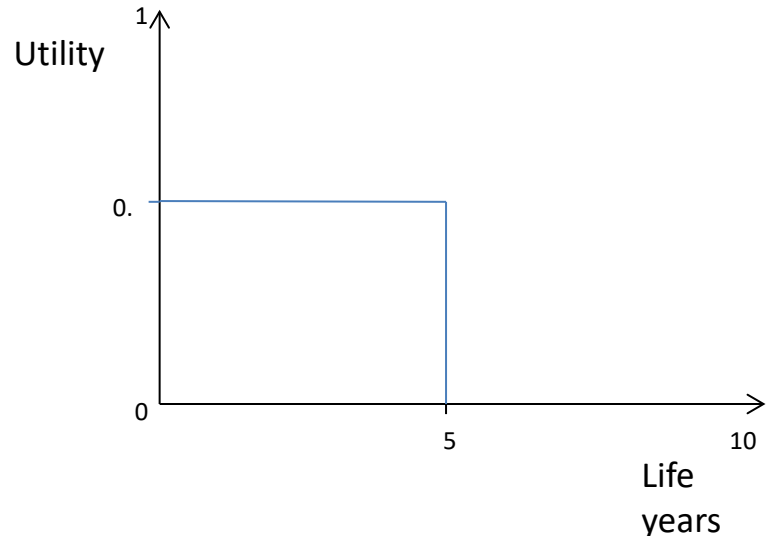
profile (Brazier et al. 2007). A disability adjusted life year (DALY) uses disability weights to adjust standard life expectancy to reflect the amount of years lost due to ill health (Brazier et al. 2007). A healthy year equivalent (HYE) considers the whole health profile as opposed to separate health states, and reflects the number of years spent in perfect health followed by instant death, which have the same utility as a profile of actual health states over an individual's expected lifetime (Kielhorn & von der Schulenburg, 2000). The discussion on CUA's within this thesis will be restricted to those which use the QALY as the measure of consequences as this is the dominant paradigm. Further details on the use of QALYs in CUA will now be provided.

2.3 Utility and the QALY

On the basis that health is a function of both the quality and length of life, the QALY aims to combine these 2 attributes into a single index figure (Brazier et al. 2007). The QALY is therefore a single measure of health outcome, which combines the morbidity and mortality of a patient. It assigns a weight to each period of time a person spends within a specific health state. This is then used to reflect a person's quality of life within one year. The weight is a single value on a scale which is anchored by 0 (a state equivalent to dead) and 1 (a state equivalent to perfect health), and where negative values can represent health states worse than dead. This value or "weight" is called a utility. A utility is a representation of an individual's preference for a health state. This is illustrated in Figure 1 below.

An individual who lives in perfect health for 5 years would have a utility value of 1. The number of QALYs over this 5 year timeframe would therefore be (1 x 5) equal to 5. However, if an individual was living in a less than perfect health state due to illness, for example 0.5, then the number of QALYs over the 5 years would be (0.5 x 5) equal to 2.5. The quality of life the patient has been in over the 5 years has therefore, been taken into account, and reflected in the number of QALYs gained.

Figure 1. The QALY.



As this measure (utility) is non-disease specific, and can be used to represent preferences across different health conditions, this approach allows comparisons to be made across populations, and across health care programmes, therefore facilitating policy decision making.

The utility value represents individuals' preference for a particular health outcome, where, the higher the utility, the greater the individuals preference for the health state. There are different methods, which can be used to calculate utilities for particular health outcomes or health states. However, all utility values should have the following key characteristics (Brazier et al. 2007):

- Utilities should be valued in a scale anchored by 0 (dead) and 1 (perfect health), where negative values reflect health states worse than dead.

- Utilities should be measured on an interval scale, where for example, the interval from 0.2 to 0.3 should represent the same to the individual as the interval from 0.8 to 0.9.

- Utilities should reflect individuals' preference for health states.

Within the literature, these weights are referred to by different terms, and can be called preferences, or values or utilities. Whilst some researchers make distinctions between these terms, within this thesis I will use the terms values or utilities.

2.4 Direct Methods of Eliciting Health State Valuations.

There are 2 components to estimating QALYs: the first is the description of the health state or health profile; the second is the valuation of these health states (Brazier et al. 1999). This section focuses on the latter: the valuation of health states. Different

methods can be used to measure either an individual's own health state, or their preferences for hypothetical health state descriptions or interventions. These include willingness to pay (WTP) and discrete choice experiments (DCE). However, traditionally, three methods are used: standard gamble (SG); time trade off (TTO); the visual analogue scale (VAS). These 3 approaches to the elicitation of utilities will now be outlined and a brief overview of their advantages and disadvantages is provided.

2.4.1 Standard Gamble

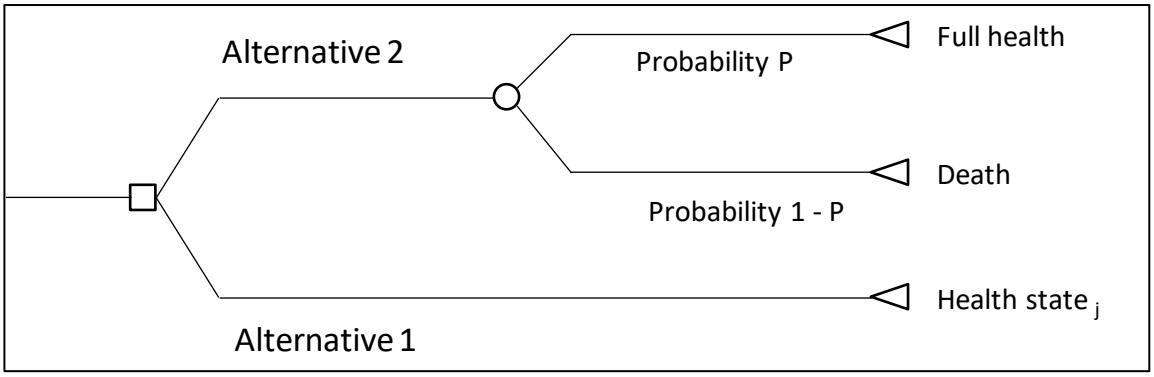
The SG methodology varies slightly depending on whether the health states being valued are chronic or temporary, and whether they are states worse or better than death (Drummond, 2005). The methodology for valuing chronic health states is described below.

The SG methodology for the valuation of health states better than death is shown in Figure 2. Individuals are presented with two alternatives: (where time = 10 years)

Alternative 1:	A certain immediate outcome (health state$_j$) for 10 years.
Alternative 2:	The uncertainty of a gamble with 2 possible outcomes: Either the patient is returned to perfect health and lives for an additional 10 years (probability *P*), or they die immediately (probability *1-P*).

The probability *P* is then varied until the respondent is indifferent between the 2 alternatives. The preference score for health state $_j$ is then equal to *P*.

Figure 2. Standard gamble for chronic health states preferred to death.



Source: Adapted from Drummond (2005).

The methodology changes for chronic health states worse than dead, and can be seen in Figure 3. Individuals are again presented with 2 alternatives: (where time = 10 years)

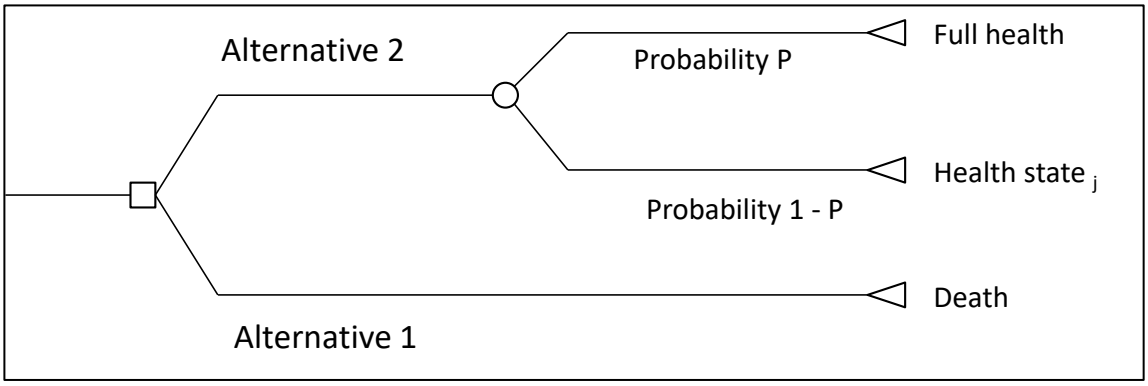
Alternative 1: A certain outcome of immediate death.

Alternative 2: The uncertainty of a gamble with 2 possible outcomes: Either the patient lives in perfect health and lives for an additional 10 years (probability *P*), or they remain in chronic health state$_j$ for 10 years (probability *1-P*).

The probability P is again varied until the respondent is indifferent between the 2 alternatives. The utility for health state$_j$ is then equal to -P(1-P). This provides a value

which ranges from minus infinity to +1, and therefore gives a greater weight to negative values (Brazier et al. 2007). Therefore, it is recommended that for health states worse than death, the values are transformed to produce a range from -1.0 to 1.0 with an equal distance from death in both positive and negative directions (Patrick et al. 1994).

Figure 3. Standard gamble for chronic health states worse than death.



Source: Adapted from Drummond (2005).

2.4.2 Time-trade off

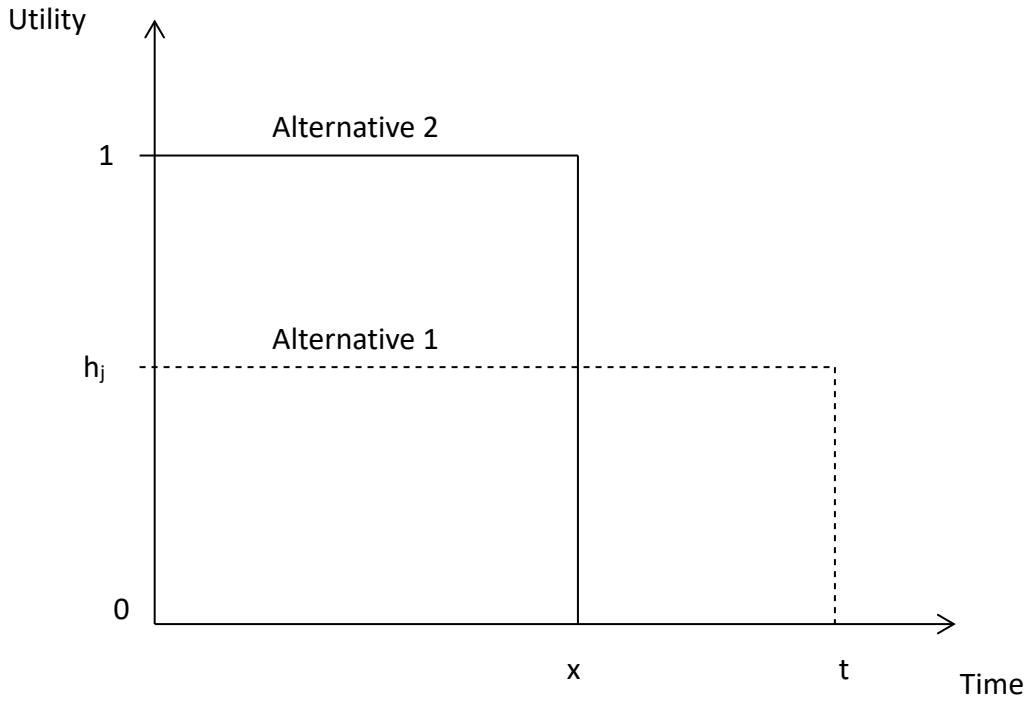
The TTO approach was developed by Torrance (1986), as an alternative to the SG. It differs from the SG in that respondents are asked to choose between two certain alternatives. There is no element of gamble included. The TTO also varies depending on whether the health states being valued are chronic or temporary, and better or worse than death. The TTO methodology for the valuation of chronic health states better than death is shown in Figure 4.

Individuals are presented with two certain alternatives:

Alternative 1:	Living for time *t* in health state $_j$ that is worse than perfect health.
Alternative 2:	Living for a shorter time *x* in perfect health.

Time *x* is then varied until the individual is indifferent between the 2 alternatives. The utility value is then equal to x/t.

Figure 4. Time trade-off for chronic health states preferred to death.



Source: Adapted from Drummond (2005).

The methodology changes for chronic health states worse than dead, and can be seen Figure 5. Individuals are again presented with two certain alternatives:

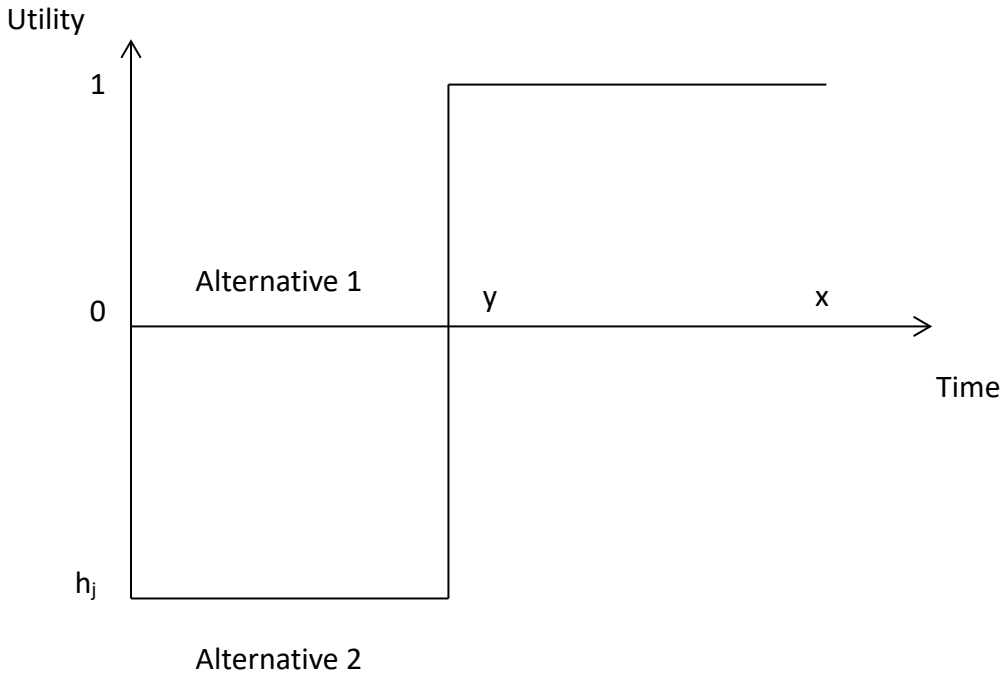
Alternative 1: Living for time *y* in health state$_j$ followed by *x* years in perfect health (where $x + y = t$) .

Alternative 2: Immediate death.

Time *x* is then varied until the individual is indifferent between the 2 alternatives. The utility value is then equal to $-x/(t-x)$ (Brazier et al. 2007). Just as for SG, for health

states considered to be worse than dead, this approach results in a value which ranges from minus infinity to +1. It is therefore recommended that the negative values are transformed onto a range from -1.0 to 1.0 (Patrick et al. 1994).

Figure 5. Time trade-off for chronic health states worse than death.



Source: Adapted from Drummond (2005).

The methodology changes for the valuation of temporary health states using TTO, as the movement from a health state to immediate death is seen as unrealistic. The approach includes 2 stages. The first can be seen in Figure 6. Individuals are presented with 2 alternatives:

Alternative 1: Living in temporary health $state_i$ (the health state of interest) for time t followed by full health.

Alternative 2: Living in temporary health $state_j$ (an anchor health state) for time x (where x<t) followed by full health.

Time x is varied until the individual is indifferent between the 2 alternatives. If the value for health $state_j$ is set to zero, then the preference for health $state_i$ = 1-x/t.

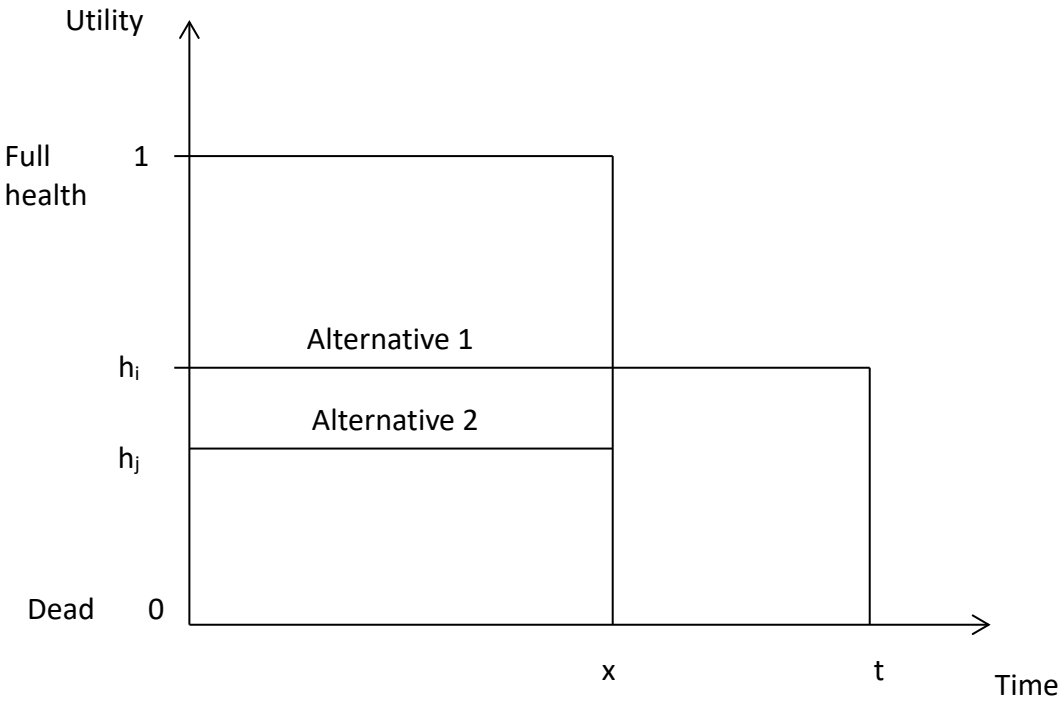
The second stage transforms this onto the 0 (dead) to 1 (full health) scale. The temporary health $state_j$ (the anchor health state) is re-defined as a short duration chronic health state. Individuals are then presented with 2 alternatives:

Alternative 1: The anchor health *j* state for time *t*, followed by immediate death.

Alternative 2: A fixed time (*x*) in full health, where *x*<*t*, followed by immediate death.

The value of health $state_j$ = x/(x-t). This is then applied to the formula in step 1 to obtain a value for h_i. This approach is used within Chapter 5 of this thesis, and will therefore be revisited.

Figure 6. Time trade-off for temporary health states.



Source: Adapted from Drummond (2005).

2.4.3 Visual analogue scale.

The VAS is much simpler to complete than the SG or TTO. It refers to a type of rating scale where respondents are shown health state descriptions, asked to rank them according to their preference, and then to mark on a scale how they would value them. They should be informed that the differences between health states should correspond with the difference in preference they have for a health state. It is argued that this results in values on an interval scale (Torrance et al. 2001. Green et al. 2000). There are many variations to how the VAS can be presented, for example, as a vertical or

horizontal line, there may be numbers on the scale, or purely descriptions of the end points (Drummond, 2005; Green et al. 2000). The health state descriptions should distinguish between temporary and chronic health states by stating the duration of the health state (Drummond, 2005). The VAS is seen as the most convenient approach to measuring preferences, as they are simple to understand and easy to administer (Drummond, 2005).

2.4.4 Other methods of preference elicitation.

There are other methods which are used to elicit preference, for example person-trade off (PTO). Here, the respondent is again asked to make a choice between two alternatives; however, these are framed as a policy decision involving groups of people. Although this approach is not commonly used, some argue it incorporates societal considerations about how resources are distributed across a population (Brazier et al. 2007). Another method increasingly used to measure health benefit is discrete choice experiments (DCE). These studies are based on the notion that health care interventions, or services can be described by their attributes, and that an individual's valuation of the intervention or service is based on the level of the attributes (Ryan, 2004).

2.4.5 Discussion

Despite the development of alternative techniques, SG, TTO and the VAS remain the 3 most commonly used methods for eliciting preferences for health states. Out of these 3, health economists recommend the use of the choice-based SG and TTO approaches (Brazier et al. 1999).

Although the most complex to complete, SG is seen by some as the "gold-standard" method of utility elicitation (Drummond, 2005). This is due to its foundation in expected utility theory (EUT) which provides a framework for how individuals make decisions under uncertainty (Von Neumann & Morgenstern, 1944). EUT states that individuals aim to maximise utility, and choose between uncertain outcomes by comparing the expected utility values of these outcomes, multiplied by the probability of receiving the outcome. The existence of both risk and uncertainty in decision making when using SG is said to reflect the risk and uncertainty experienced when receiving health care (Brazier et al. 2007). Others however dispute this, and suggest that SG results violate the axioms of expected utility as individuals responses are also influenced by factors such as risk attitude, gambling effects and loss aversion (Green et al. 2000). For example, there is the potential that a respondent is unwilling to accept any level of risk to obtain an improvement in health (Brazier et al. 1999).

The TTO was developed by Torrance et al. (1972) specifically to value health states, and its objective was to overcome the complexities of the SG technique. Studies have shown its use is associated with high response and completion rates, although these are on par with SG studies and have been exceeded by the VAS (Green et al. 2000). Although the foundation is not routed in EUT, it does have some foundations in consumer theory. Despite this, some features of TTO have been questioned, for example, TTO asks respondents to choose between 2 health states with certain outcomes. Some argue that this is not reflective of true medical decision making, where health outcomes are uncertain (Brazier et al. 2007). The TTO technique also assumes that the duration and timing of the health state has no impact on its valuation. However, empirical evidence suggests that valuations can be influenced by the duration of time spent within the health state, and an individual's preference for the timing of the health states within their lifetime (Green et al. 2000). For example, some respondents may prefer to experience an episode of ill health immediately to get it out of the way, others may prefer the reverse. These are two causes of concern for the use of the TTO technique (Brazier et al. 1999).

The VAS is seen as the simplest approach to use, and empirical evidence suggests completion and response rates are higher than those for TTO or SG (Green et al. 2000). However, there are questions in regards to its validity (Torrance et al. 2001), for example some criticise it as not being based on sound economic theory (Green et al.

2000). Although others argue that it has foundations in both economic and psychology theory (Parkin & Devlin 2006; Torrance et al. 2001). Others discuss the importance of clearly defined scale end points. Poorly defined anchors can cause problems for comparisons between respondents, and comparisons of values derived from different studies.

An additional criticism of the VAS is that it is not a choice based task, and therefore does not allow preferences to be established on a cardinal scale. This results in values which represent respondent's preferences for outcomes under certainty as opposed to utilities which reflect decisions under uncertainty and incorporating risk (Brazier et al. 1999). Parkin & Devlin (2006) question the assumption that choice based techniques are based on economic theory and "choiceless" techniques are not. They suggest this theory relates to methods for obtaining monetary values for goods and services, and not when deriving non-monetary values such as utilities.

A further criticism of the VAS method is that it is susceptible to both response spreading and context effects (Torrance et al. 2001). Response spreading refers to respondents placing their valuations for the different health states across the whole rating scale, therefore resulting in values non-reflective of their true valuation. Context effect refers to respondents' valuation of a health state being influenced by the other health states being valued (Brazier et al. 2007; Green et al. 2000). Although not the focus of this thesis, and therefore not discussed in detail, studies have been

undertaken to explore these issues further, and some have reported approaches to adjust VAS values to account for these limitations (Torrance et al. 2001; Robinson et al. 2001).

Although only reported in brief here, the debate on the strengths and limitations of direct preference elicitation methods remains. Brazier et al. (1999) and Green et al. (2000) conclude that after considering the techniques, choice-based methods remain the preferred approach to valuing health, however, Parkin & Devlin (2006) and Torrance et al. (2001) suggest the use of the VAS should not be disregarded completely.

2.5 Indirect Methods of Eliciting Health State Valuations.

The preference elicitation techniques, such as SG and TTO, previously described refer to direct methods of eliciting health state valuations, and can be complex and time consuming to complete. An alternative is the use of indirect methods. These use pre-scored multi-attribute health status classification systems, such as the EQ-5D, or the Short form 36 (SF-36) or the Short form 12 (SF-12) (Drummond, 2005). These are standardised descriptive systems, which consist of 2 components:

1. **A mechanism for describing health states:** each health state is made up of several attributes, (e.g. pain, mobility, self-care). Each attribute has a number of ordinal levels (ranging from the best e.g. no problems, to the worst e.g. severe

problems). The attributes and their levels can be combined to describe varying severities of health states.

2. **A set of values (utilities) for all of the health states that the classification system describes.** There are country specific sets of these tariffs, which have been calculated.

This method removes the necessity to complete further preference elicitation studies to obtain utility estimates (Dolan, 1997). Rather, the questionnaires can be provided to study subjects at certain study trial points, for example, pre-intervention, post intervention and at specific points of follow-up where differences in health related quality of life may be expected. The respondents can then use the questionnaires to describe their current health state. Their responses are converted into country specific utilities using the relevant tariffs. Two of the most commonly used multi-attribute preference-based measures are the EQ-5D and the SF-6D.

2.5.1 EQ-5D

The EQ-5D was developed by the Euroqol group (www.euroqol.org), and has 5 dimensions: mobility; self-care; usual activities; pain/discomfort; anxiety/depression. There are 2 versions for adults, and one version for children. The 2 adult versions differ in the number of levels included within each dimension. Within the EQ-5D-3L each dimension has 3 levels: no problem (level 1); some problem (level 2); severe problems

(level 3). By creating combinations of the domains and levels this system is able to describe 243 health states. Each health state is described using the level number of each domain (e.g. 13245). Work has been undertaken by The Measurement and Valuation Group at York University to value each of these 243 health states using the TTO technique (Dolan, 1997; MVH Group, 1995). There is therefore now a set of country-specific tariffs for each of the health states the EQ-5D system describes. More recently a second version of the EQ-5D has been developed - the EQ-5D-5L in which each of the dimensions has 5 levels: no problems (level 1), slight problems (level 2), moderate problems (level 3), severe problems (level 4), and extreme problems (level 5). The 5-level EQ-5D was developed due to concerns that the 3-level version lacked sensitivity when used in some contexts, and was designed to be more responsive to small changes in health (Devlin & Krabbe, 2013).

NICE stipulate in the 2013 Guidance on the methods of technology appraisal, that the use of the EQ-5D is their preferred measure of health-related quality of life (NICE, 2013). They recommend this should be measured directly by patients where possible, and that the valuation of the health state should be based on a valuation of public preferences from a representative sample of the UK, using a choice-based method (e.g. TTO).

2.5.2 Short-form 6D.

The SF-6D is an instrument which was originally developed from the Short form 36 (SF-36) – a health related quality of life questionnaire commonly used within clinical trials. The SF-36 consists of 36 questions related to health. Each question has a number of response choices. For example, for the physical functioning dimension the respondent can answer: limited a lot, limited a little, or not limited at all. Each response is coded, and the codes are summed to obtain dimension scores. Brazier & Roberts (2004) converted the SF-36 into a 6 dimensional health state classification called the SF-6D. They then undertook a valuation study using the SG technique on a sub-set of SF-6D health states to allow utility values to be calculated for all of the health states described by the SF-6D.

The SF-12 was developed from the SF-36 (Jenkinson et al. 1997). It provides a shorter questionnaire to measure patient's health related quality of life. Brazier & Roberts (2004) undertook further work to derive a preference based measure from the SF-12 for use in economic evaluation. They developed a preference based measure which contains 6 dimensions: Physical functioning, role limitations, social functioning, pain, mental health and vitality. Each dimension has between 3 and 5 levels. A health state is defined by selecting 1 statement from each dimension, starting with physical functioning, and ending with vitality. An example of an SF-6D health state can be seen in Figure 7.

Figure 7. SF-6D Health state 213311.

1. Your health limits you a little in moderate activities
2. You have no problems with your work of other regular daily activities as a result of your physical health or any emotional problems
3. Your health limits your social activities some of the time
4. You have pain that interferes with your normal work (both outside the home and housework) moderately
5. You feel tense or downhearted and low none of the time
6. You have a lot of energy all of the time

2.6 Process utility

As outlined within this chapter, due to finite resources and the level of demand for health care, not all services and treatments can be provided to all patients. Decisions therefore have to be made on which services and treatments investment should be directed to, and those it should be directed away from. We therefore need to make decisions on which interventions and services would provide the greatest health gain within a given budget. Within the UK we generally determine this using the QALY, which requires the collection of utility data. The standard approach to deriving utilities is to value health states, which are commonly determined using multi-attribute health classifications such as the EQ-5D or the SF-12 which include descriptions of aspects of health (e.g. pain, physical functioning) but do not contain descriptions of process. The

standard application of the utility elicitation methods therefore assumes that individuals gain no utility from the processes of receiving care, and that the value of health care is determined solely by the health outcomes it generates (Mcguire et al. 1988). This is evident when considering the SF-6D health state in Figure 6. The health which is being valued consists of a health state description, made up purely of health outcomes.

Some dispute the assumption that individuals gain utility only from the health outcomes of care (Donaldson & Shackley, 1997). It has been suggested that health economists ignore the processes that patients have to go through in the course of their treatment, and that these processes should also be incorporated into the utility function (Meenan 2001; Mooney 1998). Included in these are the "process" of treatment and additional consumption benefits such as re-assurance or information Their inclusion would take into account that the process of receiving care impacts on patients' utility, as well as the health gains they achieve (Gerard and Mooney, 1993). This utility associated with the processes of receiving care has been termed "process utility".

Although consideration of process utility may not always be relevant in an economic evaluation, in some cases it may be an important aspect of a study. For example, if the processes patients need to endure in order to receive treatment differ between trial arms. An example could be tablet versus injection administration of medication. We

may not always expect a difference in the health outcomes from these 2 modes of administration, but patients may take a tablet every day, than have to inject themselves with medication. Patients may therefore have a preference for one process over the other. This concept of process utility is not only relevant to modes of administration, but can also include levels of re-assurance, or information (Donaldson and Shackley, 1997).

The aim of this section is to introduce process utility, and to highlight that there is a debate as to whether process utility exists and whether it should be included in utility elicitation. This topic is discussed further in relation to the economic evaluation reported in Chapter 3, and subsequently forms a central component of this thesis.

2.7 Alternative ways of conducting cost-utility analyses

Within health care in the UK, CUA are commonly used to inform policy decisions on the allocation of resources. The use of the QALY as the common measure of benefits across all health care interventions allows comparisons to be made across different clinical specialities and therefore facilitates national policy decisions on resource allocation (Oliver, 2004). In order to perform a CUA, cost data, clinical outcomes and utilities are combined to obtain cost-effectiveness estimates. These different data parameters can be drawn together using different approaches. Two main approaches were described by O'Brien (1996):

2.7.1 Economic Modelling

An economic model draws together data from several different sources. For example, probabilities for clinical events may be identified from specific clinical trials, utilities may be identified from published health state valuation studies, and costs may be estimated through the application of unit costs to standard resource use within established treatment pathways. All of these data parameters are drawn together into one economic model to provide cost-effectiveness estimates. This approach O'Brien (1996) refers to as "Frankenstein's monster of economic evaluation", because it brings together many pieces of information from multiple sources and "stitches them together into a cohesive whole" (pDS100).

2.7.2 Economic Evaluations Alongside Clinical Trials

A second approach is an EEACT. Within these studies economic data are collected alongside a randomised controlled trial (RCT), the economic evaluation is therefore "piggy-backed" onto the RCT. O'Brien (1996) refers to these as Count Dracula "a vampire that feeds off the clinical trial" (pDS100). Within these piggy-back economic evaluations the economic data are collected alongside the clinical data within the same clinical trial. It is this approach which is used for the economic evaluation of ePAQ which is integral to this thesis and reported in full in Chapter 3.

Historically, the EEACT was seen as the gold standard approach to economic evaluation (O'Brien, 1996). The advantages being that the randomisation and blinding used in RCTs to reduce the potential for study bias in the hope of producing unbiased estimates of clinical efficacy or effectiveness, was thought to also translate through to the economic evaluation (Berger et al. 2003). The concurrent running of both the clinical and economic studies also allows internal consistency of the cost and outcome data retrieved, with data collected from the same group of patients, in the same setting and receiving the same treatment (Kielhorn & Schulenburg, 2000).

There are also practical advantages to this method of economic evaluation: Clinical trials are usually undertaken at an important stage in the development of a new intervention, and it is useful therefore to collect cost and resource data at this point. The concurrent running of both evaluations also enhances the collaboration between the clinicians and economists, which can only improve the standard of both evaluations undertaken (O'Brien, 1996). It has also been argued that it is more feasible financially to integrate the economic evaluation with the clinical trial rather than have a stand-alone study (O'Sullivan et al. 2005).

Although this method of economic evaluation has these advantages, there are also several disadvantages to its use. These limitations have caused much debate over the last 30 years, and this debate will now be explored further through a literature review into the limitations of EEACTs.

2.8 Literature review of limitations to EEACTs.

A literature review was undertaken to identify the debate from experts surrounding limitations to undertaking EEACTs and to highlight the current thinking on the topic. As this review is opinion-based as opposed to being data driven, the traditional methods of systematic review, which aim to retrieve, for example, well carried out RCTs of an intervention, were not appropriate methodology to use. A citation search was therefore undertaken, whereby three key articles were identified, and a search of studies, which have cited the papers, was made. These articles were selected as marking the beginning of the debate into EEACTs: O'Brien (1996) was initially identified by VB as the earliest paper to consider the strengths and weaknesses of this approach to economic evaluation. The citation search was therefore initiated from this publication. However, following discussions with supervisors it was evident that earlier work had been published by both Drummond & Stoddart (1984) and Drummond & Davies (1991). Therefore these 2 additional papers were subsequently included, resulting in 3 core articles.

The review was therefore initiated from the following 3 key texts:

Drummond & Stoddart (1984) marks the beginning of the debate on the collection of economic data within clinical trials. When written, the number of economic evaluations being undertaken in health care was increasing, and there was a growing

demand for incorporating economic analysis into clinical trials in order to obtain timely economic data based on sound clinical evidence. Drummond & Stoddart (1984) put forward potential objections, which could be raised by a research team concerning the inclusion of economic end points and data collection within a clinical trial. These concluded that the addition of an economic component adds to both the complexity and the cost of the clinical trial, and there is often a lack of economic expertise within a clinical study team to provide advice on the study design. They also highlighted that commonly clinical trials are designed to evaluate the clinical efficacy of a pharmaceutical product, and are therefore often compared to a placebo – it is unlikely this is the correct comparator for a cost-effectiveness analysis, where the key comparison is with the current, standard clinical care. They also suggested that, even where the clinical trials have suitable comparators for an economic evaluation, the experimental nature of clinical trials creates an environment which is not reflective of routine clinical care.

Drummond & Stoddart (1984) propose certain arguments which should be considered when deciding whether to include economic analyses within clinical trials. These include, considering whether the cost-effectiveness outcomes are required to inform an imminent resource allocation decision, or whether there are likely to be large resource consequences from the introduction of a drug or intervention to the market – for example, if considerable differences in unit costs are expected between standard

care and a new intervention. Their paper also discusses key data requirements for economic evaluations. These include measures of both costs and benefits to patients, inpatient and community costs, the inclusion of productivity costs, and the collection of utility data which can be converted into QALYs. They conclude by suggesting a "phased" approach to the consideration and implementation of economic evaluation alongside the clinical trial phases throughout the product development cycle. For example, within Stage 1 making a rough estimate of the costs associated with comparators within the explanatory phase, and using this to decide whether future economic work is required or warranted. If economic data is desired Stage 2 requires the identification and selection of an appropriate clinical trial and finally, in Stage 3 they suggest awaiting initial clinical trial outcomes to determine whether detailed economic data collection are required. They suggest the implementation of this 3 staged approach would reduce the burden on the trial team, and allow consideration to be made on potential policy implications once the trial results are known.

Drummond and Davies (1991) revisited these methodological issues in a subsequent paper. They discuss the clashing purposes of clinical and economic trials, and important issues which should be considered. For example, the need to choose therapies which reflect current clinical practice, to obtain a sample size large enough to identify a difference in costs, to select a location which could be generalised more easily in other areas.

The third paper is O'Brien's (1996), which uses the "Frankenstein" and "Count Dracula" analogy previously discussed, to differentiate between the 2 types of economic evaluations. He goes on to highlight problems associated with the latter: EEACTs. He refers to "seven threats to the validity" of these trial based economic analyses. These are reported and discussed within the literature review results.

2.8.1 Literature review methodology.

Citation searches were initiated using the 3 key texts outlined above. The searches were initially performed in 2008, and were re-run in July 2015. They were undertaken using the following databases:

- Medline via Ovid
- Cochrane library (a methodological search using the term "Economic evaluation alongside a clinical trial")
- Web of Science (formerly Web of Knowledge).

Inclusion Criteria

- Studies containing comment pertaining to methodological issues of undertaking EEACTs.
- English language

Exclusion Criteria

- Studies containing no relevant information
- Economic evaluations
- Conference abstracts
- Non-English papers

Inclusion/Exclusion Process

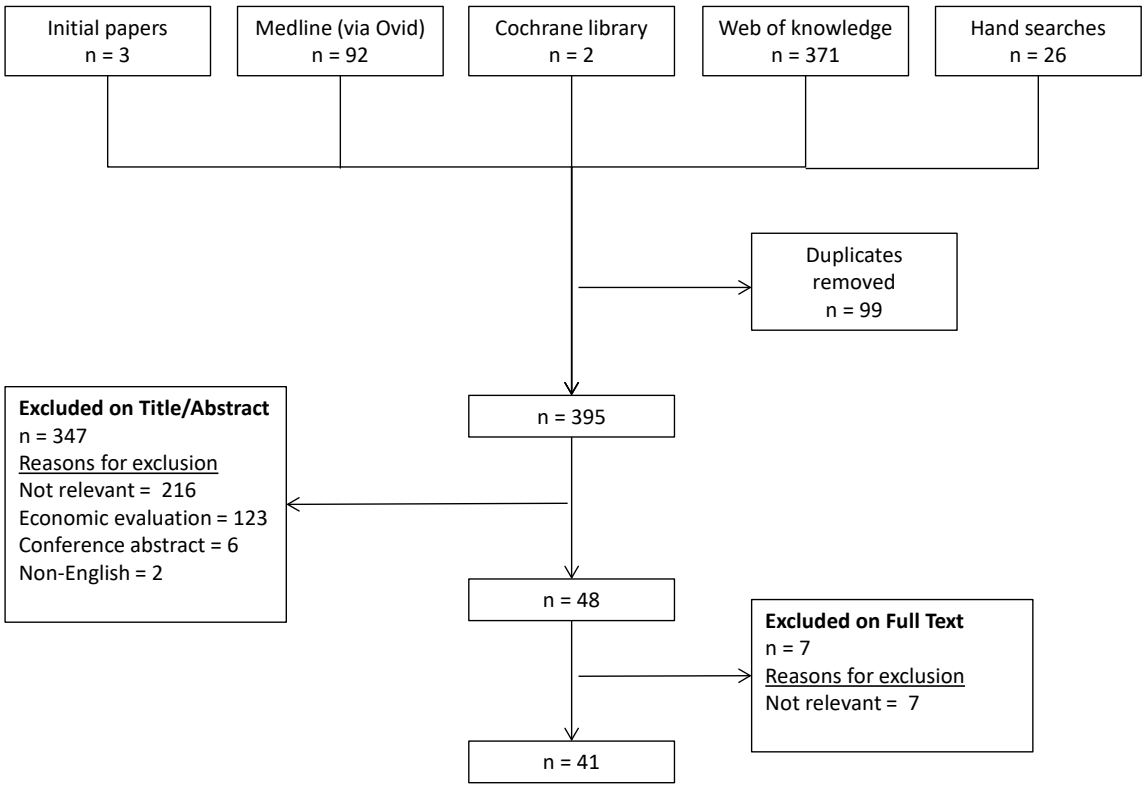
A list of all studies cited by the 3 key papers was obtained. Following the removal of duplicates, the titles and abstracts of the remaining studies were reviewed against the inclusion/exclusion criteria. Those which could be excluded using the title and abstract information were. The full text documents of the remaining studies were obtained, and again reviewed against the inclusion/exclusion criteria for inclusion in the review. The reference lists of all included studies were also reviewed for relevant papers.

2.8.2 Literature review results.

The initial searches identified 494 titles. 99 duplicates were removed. Following a review of the titles and abstracts of the remaining studies, 347 studies were excluded. Full text documents for the remaining 48 studies were obtained and reviewed for inclusion. Seven of these were excluded. 41 studies which met the inclusion criteria

were identified and included in the review. Publication dates ranged from 1984 to 2015. A PRISMA diagram showing the number of studies included and excluded at each stage, and reasons for exclusion can be seen in Figure 8. A list of included studies can be seen in Appendix 1.

Figure 8. PRISMA Diagram.



The Threats to the Validity of EEACTs summarises the threats to the validity of economic evaluations discussed within the included papers, and provides an overview of each. Figure 9 shows the number of papers which discussed each threat.

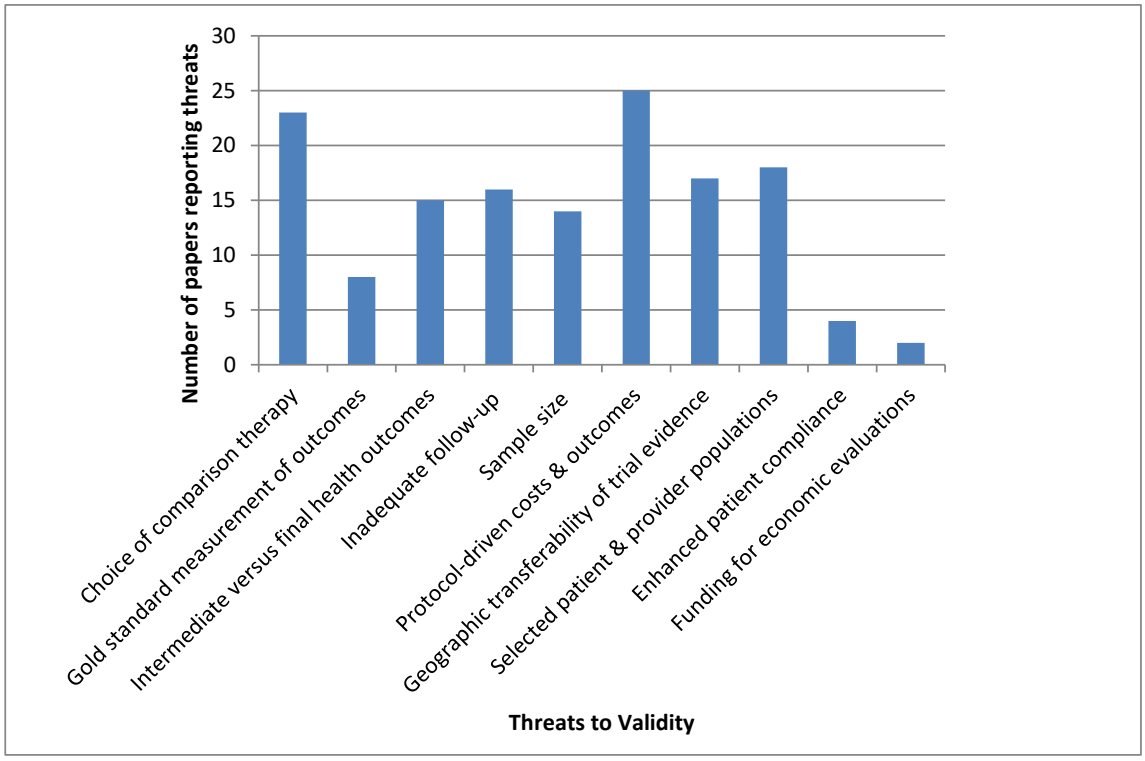
Table 1. The Threats to the Validity of EEACTs

Threat to Validity	Description
Choice of comparison therapy	With a clinical trial aiming to measure efficacy, the comparator is often a placebo. Within economic evaluations to inform policy decisions, the comparators should be standard practice (Sheldon, 1996).
Gold standard measurement of outcomes	Within the trial setting, the monitoring patients receive is often more stringent than those patients receiving routine clinical care. This may include more or different tests, and more regular visits to the trial centre. This increased monitoring will increase the probability of finding new side effects and diseases, which may not have been found through routine monitoring (Ramsey et al. 2015).
Intermediate versus final health outcomes	Within clinical trials, it is usual to collect intermediate health outcomes, (for example, blood pressure; cholesterol levels), as opposed to final health outcomes (e.g. mortality). This is due to practicalities such as realistic trial durations. However, for economic evaluations, the final health outcomes are important (Cookson & Hutton, 2003).
Inadequate follow-up	The time-frame for a clinical trial is usually based on the time needed to capture clinically meaningful differences in relevant treatment outcomes. In contrast, economic evaluations usually require life-time morbidity and mortality data (Sculpher et al. 2006).

Threat to Validity	Description
	In addition, within a clinical trial, patients who experience an adverse event are usually then excluded from the study. Within an economic evaluation, these patients should be followed and their long term costs and outcomes should be included in the analysis.
Sample size	The sample size within a trial is usually powered to detect a clinically meaningful difference in the primary clinical outcome. This is usually not a big enough sample to detect a statistically significant difference in the costs, particularly as costs tend to vary considerably between patients (Ramsey et al. 2015).
Protocol-driven costs and outcomes	Within a clinical trial patients often have tests and hospital visits, above those which would be received in standard clinical care. The strict monitoring and research conditions introduced by a RCT may represent a misleading impression of the outcomes of routine care (Marshall & Hux, 2009). Protocol led monitoring of patients could also lead to "case-finding", whereby the protocol driven tests identify previously undetected conditions.
Geographic transferability of trial evidence	Economic evaluations are primarily undertaken to allow decision makers to come to an informed decision regarding the allocation of health care resources (Kielhorn & Schulenburg, 2000). The information must therefore be relevant to specific decision makers' needs, and reflect their specific population. As the debate has already indicated,

Threat to Validity	Description
	the outcomes and costs associated with an RCT are for many reasons not representative of standard practice but rather are pertinent only to the trial population.
Selected patient and provider populations	Within an RCT there are stringent inclusion and exclusion criteria, leading to the exclusion of certain co-morbidities or age groups. Although this can enhance the clinical outcomes of a study, it introduces a further limitation to an economic evaluation alongside a clinical trial as it is unlikely the results are generalisable to the wider population (O'Sullivan et al. 2005).
Enhanced patient compliance	Some suggest that compliance with treatment is higher within clinical trials than standard care. This is partly due to the strict monitoring inherent in the trial, but also by the nature of the fact that the participants have agreed to participate. Therefore, when applying study results to the general population, it should be expected that compliance levels are lower. (Bombardier & Maetzel, 1999)
Funding for economic evaluations	Studies have shown that the source of funding for economic evaluations can impact on the reporting of results. Friedberg et al. (1999) for example, explored the impact of the funding source on outcomes, and concluded that economic evaluations sponsored by pharmaceutical companies were less likely to report unfavourable outcomes (Friedberg et al. 1999).

Figure 9. Bar Chart showing the number of papers reporting each threat to the validity of EEACTs



2.8.3 Discussion

This literature review has highlighted the debate on the limitations to EEACTs. A key conflict noted by all studies is the fundamental difference in the research questions and objectives between an RCT and an economic evaluation, and the conflicting fight for internal versus external validity. RCTs are specifically designed to measure safety and efficacy, and are therefore undertaken within controlled environments, usually with strict inclusion/exclusion criteria for patients and close clinical monitoring. This enhances the internal validity of the trial. Economic evaluations, however, aim to

obtain a measure of cost-effectiveness within routine clinical care. The characteristics which therefore enhance the internal validity of the clinical trial reduce the external validity required for undertaking meaningful economic evaluations.

These conflicting objectives significantly influence the trial design, and lead to many of the limitations identified. For example, the comparator of choice within an RCT is usually a placebo, therefore allowing for the therapeutic efficacy of the treatment to be identified, separate from any placebo effect related to its delivery. However, the comparator (or comparators) of choice for an economic evaluation, which is designed to ascertain whether a new intervention is more cost-effective within standard care, is current routine care (Drummond & Stoddart, 1984; Marshall & Hux, 2009), and ideally comparisons should be made with all available treatment options allowing all relevant information to be considered in decision making processes (Sculpher, 2015).

The most commonly discussed limitation is the importance of recognising and excluding protocol-driven resource use and outcomes. These refer to those which have been introduced to the study due to undertaking the research, rather than those actually linked to the intervention (O'Brien, 1996). The strict research conditions introduced by an RCT may give a misleading impression of resource use and outcomes in routine practice (Cookson & Hutton, 2003). RCTs rarely provide insight into the full range of treatment strategies which may be possible for an intervention, for example alternative treatment sequences or the use of starting and stopping rules (Sculpher

2015). Within clinical trials, there are often specific guidelines regarding the frequency and intensity with which visits, tests and other health care services should be provided, and these often differ from the processes which occur within standard practice, and therefore provide resource use estimates non-reflective of current clinical care (O'Sullivan et al. 2005). This increased monitoring may also identify outcomes which would not be picked up within routine care. Using the resource use and outcomes directly from the clinical trial therefore, without accounting for the deviation away from routine care is unlikely to give a true representation of the cost-effectiveness of an intervention in practice.

There are further limitations linked to the measurement of outcomes collected and required for the clinical and economic studies. One of these relates to the often detailed and expensive tests and procedures which are undertaken within an RCT compared to the less stringent, equivalent tests in standard care (O'Brien, 1996). Not only will this lead to higher resource use estimates than used in routine care, but if the tests are more stringent and comprehensive within an RCT it is also likely that more side effects or adverse events will be picked up, and at an earlier stage. The repeated testing and regular visits to the trial centre will also increase the probability of finding new side effects and diseases which would not have been found in routine care (Ramsey et al. 2001).

There is also likely to be a clash between the outcomes required for clinical and economic analyses. Clinical trials often use composite end points, where multiple clinical endpoints are combined, for example including both fatal and non-fatal events, or cost per cardiovascular event (Ramsey et al. 2015). This has the advantage of reducing the sample size required to detect significant differences between primary endpoints, allowing for smaller, less costly trials to be run. The use of these composite endpoints in economic evaluations however is not ideal, as the cost per composite endpoint is unlikely to be useful. Rather, the outcomes of interest would be costs per single outcome, for example, cost per myocardial infarction or stroke.

The ideal timeframe for the collection of clinical outcomes also differs between the study types. Clinical endpoints within trials usually collect data on intermediate outcomes e.g. blood pressure or cholesterol levels, whereas the ideal outcomes for the economic evaluations are lifetime endpoints e.g. life years gained. This is primarily due to the practicalities of undertaking a trial for long enough to generate data on long term outcomes, plus the sensibility of undertaking a lifetime study and the usefulness of the data it would generate. It is therefore accepted that the data collected in an EEACT may need to be supplemented with additional long term data to extrapolate outcomes to a lifetime horizon.

There may therefore be a lack of patient level data available on long term outcomes due to short trial durations, but in addition to this it is usual within a clinical trial that if

a patient experiences a severe adverse event e.g. a stroke or myocardial infarction, they are then excluded from any further involvement in the trial. However, in reality, and for the economic evaluation, these patients should continue to be followed and the costs and outcomes for their adverse event should also be included in the analysis.

Further limitations, and potential biases are linked to the population and sample size. Several papers discussed the stringent inclusion and exclusion criteria applied in clinical trials, which leads to the exclusion of patients with certain co-morbidities or age groups (Sculpher 2015). Although this can enhance the clinical outcomes, it introduces a further limitation in EEACTs, as it creates a select study population which is not representative of those obtaining routine care, and therefore, it is unlikely the clinical trial results will be generalisable to the wider population (O'Sullivan et al. 2005).

It is also evident that, even where economic data are being collected, the sample size is often determined on power calculations to ensure a significant difference in the primary clinical endpoint is detectable. In reality the sample size required to detect a significant difference in costs is much larger, due to the increased variability of cost outcomes. Ramsey et al. (2015) do however suggest that new techniques are being developed to overcome this limitation, based on the expected value of perfect information methods.

There are also some more subtle points to consider. As highlighted by Rittenhouse (1996), those patients who agree to participate in research are often the more compliant patients and are often healthier patients. It can also be argued that patients who accept to be randomised to a treatment group tend to have a preference for the new treatment. This therefore biases the outcome in favour of the new treatment as their preferences reinforce their self-assessed outcomes. This introduces an additional component of trial design – patient preferences, and will be discussed further within this chapter, specifically related to the ePAQ RCT.

Just as patients included within trials are likely to differ from those receiving standard care, so too are the trial centres. Often, clinical trials are undertaken within university hospitals, and patients tend to see consultants (whose time is more expensive), as opposed to more junior staff. As stated by Rittenhouse (1996) "The provider population in most trials is far from representative of the general provider population." (p.24). This again, is likely to result in cost estimates which are not truly reflective of routine clinical care.

A further point raised by several of the studies is the requirement for utility data to be incorporated into the economic evaluation. Utility estimates can be obtained using several methods, the most common of which is the inclusion of a generic, preference based quality of life measure such as the EQ-5D or SF-6D. The results from these measures can be combined with national tariffs to obtain utility values reflecting the

general population's preference for the health outcomes (Bonsel et al. 1993). These utility values are then used to calculate QALYs and for determining the incremental cost effectiveness ratios (ICERs) of new interventions.

The most recent paper identified (Ramsey et al. 2015) discusses the advantages of using current technology to assist in the collection of cost-effectiveness data. For example, tracking health status and resource use using smart phones.

2.8.4 Conclusion

This literature review has shown that the debate around the limitations to EEACTs continues. Early papers such as O'Brien et al. (1996) discuss the balance between obtaining the internal validity imposed by the controlled trial environment, and external validity required for an economic evaluation, and this discussion is continued to the present day (Ramsey et al. 2015), with Sculpher (2015) identifying the limitations which could be addressed using modelling e.g. the use of an absolute measure of benefit such as a hazard ratio, being applied to a baseline measure in a model, as opposed to using relative effectiveness from an RCT; or the extrapolation of outcomes beyond the RCT trial horizon. Despite their limitations, the use of EEACTs to determine cost-effectiveness still fuels this active debate, and therefore to some RCTs are still seen as an opportunity for the collection of patient level resource use data, and

methods of undertaking EEACTs continue to evolve and improve over time (Ramsey et al. 2015).

Some argue that the limitations are so evident because importance is usually placed on the clinical, as opposed to the economic outcomes even in the trial planning stage (Marshall & Hux, 2009). Therefore, trials are primarily designed to obtain data to show clinical efficacy, and the economic trial design and outcomes are seen as secondary. Several papers emphasise that in order to reduce these limitations, health economists should be included early on in the trial planning stage, as opposed to being brought in following the development of the clinical protocol, and requested to add-in economic endpoints as an adjunct to the clinical trial (Cook et al. 2004; Briggs, 2000; Drummond & Stoddart, 1984).

Within the next chapter, the traditional CUA framework is applied in an EEACT into the cost-effectiveness of an electronic pelvic floor assessment questionnaire (ePAQ) used in conjunction with a telephone consultation, compared to standard care. The discussion relates the findings to the threats to the validity of EEACTs identified within this review.

2.9 References

- Berger, ML. Gingefors, K. Hedblom, EC. Pashos, CL. Torrance, GW. Smith, MD. (Eds.) (2003) *ISPOR book of terms: Health care cost, quality and outcomes*. New Jersey, ISPOR.

- Bombardier, C; Maetzel, A. (1999) Pharmacoeconomic evaluation of new treatments: efficacy versus effectiveness studies? *Ann Rheum Dis*, 58(Suppl 1), I82-5.

- Bonsel, GJ. Rutten, FF. Uyl-de Groot, CA. (1993) Economic evaluation alongside cancer trials: methodological and practical aspects. *Eur J Cancer*, 29A(Suppl 7), S10-4.

- Brazier, JE. Roberts, J. (2004) The Estimation of a preference-based measure of health from the SF-12. *Medical Care,* 499, 851-859.

- Brazier, J. Deverill, M. Green, C. Harper, R. Booth, A. (1999) A review of the use of health status measures in economic evaluation. *Health Technology Assessment*, 3(9),1-164.

- Brazier, J. Ratcliffe, J. Salomon, JA. Tsuchiya, A. (2007) *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford, Oxford University Press.

- Briggs, A. (2000) Economic evaluation and clinical trials: size matters. The need for greater power in cost analyses poses an ethical dilemma. *BMJ*, 321(7273),1362-1363.

- Cook, J. Drummond, M. Heyse, JF. (2004) Economic endpoints in clinical trials. *Stat Methods Med Res*, 13(2), 157-76.

- Cookson, R. and Hutton, J. (2003) Regulating the economic evaluation of pharmaceuticals and medical devices: a European perspective. *Health Policy,* 63,167-178.

- Cookson R. (2003) Willingness to pay methods in health care: a sceptical view. *Health Economics*, 12, 891-894.

- Devlin, NJ. and Krabbe, PF. (2013) The development of new research methods for the valuation of EQ-5D-5L. *Eur J Health Econ,* 14(Suppl 1),S1-S3.

- Dolan, P. (1997) Modelling valuation for EuroQol health states. *Med Care* 35, 351–363.

- Donaldson, C. and Shackley, P. (1997) Does "process utility" exist? A case study of willingness to pay for laproscopic cholcystectomy. *Soc Sci Med*, 44, 699-707

- Drummond, MF. and Stoddart, GL. (1984) Economic analysis and clinical trials. *Controlled clinical trials* 5,115-128.

- Drummond, MF. and Davies, L. (1991) Economic analysis alongside clinical trials. Revisiting the methodological issues. *Int J Technol Assess Health Care*, 7(4),561-73.

- Drummond, MF. Schulpher, MJ. Torrance, GW. O'Brien, B. Stoddart, GJ. (2005) *Methods for the Economic Evaluation of Health Care Programmes*. 3rd Edition, Oxford, Oxford Press.

- Friedberg, M. Saffran, B. Stinson, TJ. et al. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. JAMA 1999; 282(15).

- Gerard, K. and Mooney, G. (1993) QALY League Tables: handle with care. *Health Economics*, 2, 59-64.

- Green, C. Brazier, J. Deverill, M. (2000) Valuing Health-related quality of life. A review of Health State Valuation Techniques. *Pharmacoeconomics*, 17(2), 151-165.

- Jenkinson, C. Layte, R. Jenkinson, D. Lawrence, K. Petersen, S. Paice, C. Stradling, J. (1997) A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *Journal of Public Health Medicine,* 19(2), 179-186.

- Kielhorn, A. and Graf von der Schulenburg, JM. (2000) *The health economics handbook.* 2nd Edition, Chester, Adis International.

- Marshall, DA. Hux, M. (2009) Design and analysis issues for economic analysis alongside clinical trials. *Med Care*, 47(7 Suppl 1), S14-20.

- Mcguire, A. Henderson, J. Mooney, G. (1988) *The Economics of Health Care. An Introductory Text*. London, Routledge Publishers.

- Meenan, R. (2001) Developing appropriate measures of the benefits of complementary and alternative medicine. *Journal of Health Services Research and Policy*, 6(1), 38-43.

- Mooney, G. (1998) The benefits of health care. Health Care Analysis, 6(2),99-101.

- MVH Group. (1995) The Measurement and Valuation of Health. Final report on the modelling of valuation tariffs. York Centre for Health Economics.

- NICE (2013) National Institute for Health and Care Excellence. Process and Methods Guides: Guides to the methods of technology appraisal 2013. Available from: http://publications.nice.org.uk/pmg9 [Accessed 28th August 2015).

- O'Brien, B. (1996) Economic Evaluation of Pharmaceuticals. Frankenstein's monster or Vampire of trials? *Medical Care*. 34(12)DS99-DS108 Supplement.

- O'Sullivan, A. Thompson, D. Drummond, M. (2005) Collection of health economic data alongside clinical trials: is there a future for piggyback evaluations? *Value in Health* 8(1) 67-79.

- Parkin, D, and Devlin, N. (2006) Is there a case for using visual analogue scale valuations in cost-utility analysis. *Health Economics*, 15, 653-664.

- Patrick, DL. Starks, HE. Cain, KC. Uhlmann, RF. Pearlman ,RA. (1994) Measuring Preferences for Health States Worse than Death. *Medical Decision Making,* 14,9-18.

- Ramsey, SD. McIntosh, M. Sullivan, SD. (2001) Design issues for conducting cost-effectiveness analyses alongside clinical trials. *Annu Rev Public Health*. 22,129-41

- Ramsey, SD. Willke, RJ. Glick, H. Reed, SD. Augustovski, F. Jonsson, B. et al. (2015) Cost-effectiveness analysis alongside clinical trials II-An ISPOR Good Research Practices Task Force report. *Value Health*, 18(2),161-72.

- Rittenhouse, B. (1996) *Uses of models in economic evaluations of medicines and other health technologies.* Office of Health Economics, BSC Print, London.

- Robinson, A. Gyrd-Hansen, D. Bakon, P. Baker, R. Pennington, M. Donaldson, C. EuroVaQ Team (2013) Estimating a WTP-based value of a QALY: the "chained" approach. *Soc Sci Med*,92, 92-104.

- Robinson, A. Loomes, G. Jones-Lee, M. (2001) Visual analog scales, standard gambles, and relative risk aversion. *Med Decis Making,* 21, 17-27.

- Ryan M (2004) Discrete choice experiments in health care. *BMJ,* 328,360-1.

- Sculpher, M. (2015) Clinical Trials Provide Essential Evidence, but Rarely Offer a Vehicle for Cost-Effectiveness Analysis. *Value in Health,* 18,141-142.

- Sculpher, MJ. Claxton, K. Drummond, M, McCabe, C. (2006) Whither trial-based economic evaluation for health care decision making? *Health Economics*, 15(7), 677-687.

- Sheldon, T. (1996) Problems of using modelling in the economic evaluation of health care. *Health Economics* 5, 1-11.

- Torrance, GW. (1986) Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5, 1-30.
- Torrance, GW. Feeney, D. Furlong, W. (2001) Visual Analog Scales: Do they have a role in the measurement of preferences for Health States? *Med Decis Making,* 21,329-334.

Chapter 3: An economic analysis alongside a clinical trial in pelvic floor medicine.

3.1 Introduction

This chapter applies the standard methods of economic evaluation for cost-utility analysis. It reports an EEACT which was organised by Mr Stephen Radley (Consultant Gynaecologist), and was developed and performed through collaboration between Sheffield School of Health and Related Research (ScHARR) and Sheffield Teaching Hospitals (STH). The economic analysis was developed by Professor Simon Dixon, and the clinical analysis by Dr Georgina Jones. The research was funded by a charitable grant from Sheffield Hospitals Charitable Trust. The EEACT aimed to assess the cost-effectiveness of an online assessment tool used in conjunction with a telephone consultation, compared to standard care.

The structure of this chapter is firstly to introduce and describe ePAQ. The rationale for the EEACT is explained, and the study methods and results are reported. Finally, the chapter discusses the ePAQ trial in relation to the limitations of EEACTs identified in the literature review in Chapter 2. Subsequent chapters aim to explore one limitation further.

3.2 Background

ePAQ is a web-based, interactive questionnaire (electronic pelvic floor assessment questionnaire), which provides a detailed evaluation of a woman's pelvic floor symptoms and their impact on her quality of life (http://www.epaq.co.uk/), and can be completed online or at clinic via a touch screen or computer. The questionnaire is divided into four dimensions: urinary, bowel, vaginal and sexual.

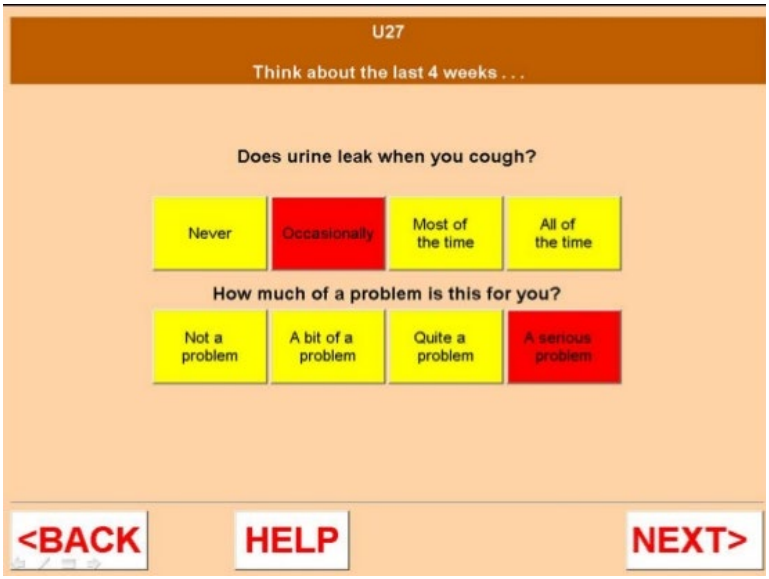
Each dimension has a number of items (urinary 35, bowel 33, vaginal 22 and sexual 28). Within these 4 dimensions there are 19 scored domains. All items that contribute to the domains score between 0 and 3 (0 indicating worst health, and 3 indicating best health). This is converted into a score ranging from 0 to 100, where 0 indicates the best health state and 100 the worst. If there are no clinically meaningful findings the score is reported as "screen negative" (Jones et al. 2008).

The questionnaire asks patients about specific symptoms. When they report the presence of a specific symptom they are automatically presented with impact questions relating to how "bothersome" they find that symptom. These aim to combine clinical findings with the impact these symptoms have on the patient. This can be seen in the screen shot in Figure 10. These impact questions do not appear if the patient responds "no" to suffering a specific symptom.

On completion of the questionnaire the patient receives a results summary. This contains results for each of the 4 dimensions and provides scores for the symptoms within each domain and scores reflecting the impact these symptoms have on the patient's life. The impact scores are presented as a clock face where an empty circle represents "not a problem", 1/3 circle represents "a bit of a problem", 2/3 circle represents "quite a problem" and a complete circle represents "a serious problem". Patients can click on this impact score and obtain more information relevant to that domain. A more detailed report on each specific domain can also be printed from the responses reporting specific results for each question answered (www.epaq-online.co.uk).

ePAQ was developed with the aim of creating an assessment tool for symptoms and quality of life related to bowel, bladder, vaginal and sexual function in women. The aim of the tool is to provide a comprehensive summary of a patient's present condition which can be used to aid in the clinical process (Jones et al. 2008).

Figure 10. Example of an ePAQ bothersome question.



3.2.1 The Rationale for the ePAQ EEACT.

The ePAQ-pelvic floor questionnaire is currently administered to all patients attending the urogynaecology department in advance of their clinic appointments. It is then used by the consultant as an adjuvant to the patient's consultation. The development of an online version of ePAQ means patients are able to access and complete the questionnaire at home, in advance of their clinic appointment, which allows the information to be used in the triaging of patients to the most appropriate clinics in both primary and secondary care.

The ability of patients to complete the tool without attending the hospital has also facilitated the introduction of telephone consultations in place of face to face clinic visits. The consultant is now able to obtain a clear and concise summary of a patients

signs and symptoms, and the impact these have on her life, without having to undertake a prolonged consultation, which, due to the nature of the disease area, may include questions which women find embarrassing to answer. The information is supplemented by a telephone consultation which allows the consultant to obtain further details on the patient's condition and the triaging of patients to appropriate services if required.

It was believed that the use of ePAQ online in conjunction with telephone consultations, and the move away from traditional face to face consultations, could potentially have significant economic benefits, whilst also enhancing the patients' experience of care. An economic evaluation was therefore undertaken. An EEACT to compare the impact of a virtual clinic with standard care was therefore developed. The initial research team included clinicians, clinical researchers and health economists at STH and The School of Health and Related Research (ScHARR) at the University of Sheffield. VB joined the team at the start of patient recruitment.

3.3 Methods

An EEACT was undertaken which aimed to assess the impact of patients using ePAQ in advance of clinic appointments, combined with a telephone consultation, on patient care. The study protocol was approved by the North Sheffield ethics committee and is registered at www.clinicaltrials.gov (identifier NCT02176330).

Both the clinical study design and methods for the economic evaluation which ran alongside the trial were initially devised by members of ScHARR in conjunction with STH staff. A subsequent amendment to the ethics application was facilitated by V.B, and is described within this Chapter.

The clinical study analysis was performed by G.J and although not reported in depth here, the clinical study results are included in appendix 5. This paper has currently been submitted for publication. This chapter reports the methodology and results of the cost-effectiveness analysis. The economic analysis was performed by V.B.

3.3.1 Study Sample

All women who had been referred to the STH urogynaecology services, aged 18 years or more, and who were able to read and understand the English language were invited to enter the study. Over a 12 month period beginning in March 2007 potential participants were identified through a review of the referral letters received from GPs and other clinicians by the consultant to whom the patient had been referred.

Eligible patients were telephoned to discuss the study with a research nurse. Interested patients were sent an information sheet and consent form. Patients who did not return the signed consent form within 1 week were telephoned again to ask whether they were still interested in participating. Patients who were not contactable or who declined to participate were sent an appointment for usual care in the clinic.

On receipt of signed consent forms, patients were randomised to one of two groups:

- **Intervention:** Patients completed ePAQ online remotely and subsequently received a telephone consultation;
- **Control:** Patients completed ePAQ online remotely or in clinic pre-consultation. They then attended a standard face-to-face outpatient clinic appointment.

The sample size calculation was performed by G.J, and the study was powered to detect a difference in the primary clinical outcome: the Patient Experience Questionnaire (PEQ). The study was not powered to detect differences in cost-effectiveness outcomes.

3.3.2 Intervention Group

All women randomised to the intervention group were entered onto the hospital PAS system under the ePAQ clinic list. They were then posted an information letter which included instructions on how to access and log-on to the ePAQ online system. On completion of the questionnaire patients could print out a summary of their questionnaire data if desired. They were then asked to forward their username and questionnaire ID to a STH email address (epaq@sth.nhs.uk). In line with good data protection practice, all data were encrypted and anonymised, and only accessible by an approved clinician with access to the questionnaire ID and username forwarded by the

patient. The clinician used the patient's ePAQ data alongside their clinical case notes and the original referral letter as an adjunct to the subsequent telephone consultation.

The research nurse then arranged the patient's telephone consultation. Those who were unable to complete the ePAQ questionnaire online at home could use this telephone call to arrange to attend and complete the questionnaire at the hospital.

3.3.3 Control

All women randomised to the control group were posted an appointment to attend the urogynaecology clinic as is standard in routine care at STH. They were given the choice to complete ePAQ either online at home in advance of the appointment, or to complete on arrival at the clinic in advance of their appointment.

3.3.4 Outcome measures

The following section outlines the outcome measures used in the cost-effectiveness analysis. The clinical analysis was completed by G.J and is currently being written up for publication. Clinical outcomes included the patient experience questionnaire (PEQ), client satisfaction questionnaire (CSQ-8) and the QQ-10 which is designed to assess patients' views on questionnaire use in health care.

PEQ

The PEQ is a brief questionnaire which was developed in primary health care and is used for measuring patients' experience of interaction, emotion and consultation outcome (Steine et al. 2001). The questionnaire consists of 18 items on 5 dimensions: communication; emotions; short term outcome; barriers; and relations with auxiliary staff. Scoring involves sum scales for each of the 5 dimensions. Four of the scales run from 1 to 5 and the emotion scale runs from 1 to 7. A high score represents a good communication experience, positive emotions, positive consultation outcome, no communication barriers and good relations with the staff. For the ePAQ RCT only 4 dimensions were included. The respondents' relations with auxiliary staff were deemed irrelevant and were therefore omitted.

CSQ-8

The CSQ-8 is a questionnaire which is used to measure satisfaction with services received by individuals and families (Larsen et al. 2007). Scoring of the CSQ involves an unweighted summation of the direction-corrected response values and a calculation of measures of central tendency of the individual item ratings and for the total scale scores.

3.3.5 Utilities.

In order to obtain the utility data required to calculate the number of QALYs gained, patients needed to complete the SF-12 at both baseline and 6 months follow-up. As detailed in Chapter 2, the SF-12 is a standardised, multidimensional, generic measure of health related quality of life (HRQOL), which can be converted into utilities using the SF-6D (Brazier & Roberts, 2004). These utility values, which are anchored on a scale of 0 (death) to 1 (full health), represent societal preferences for a health condition, and are combined with the duration of the health state to calculate quality adjusted life years (QALY).

Following study initiation, it was noted that despite the study protocol stating a cost per QALY as a study outcome, the SF-12 had not been included in the original study protocol or gained ethics approval, and had not been sent out to patients. V.B therefore submitted an amendment to the study protocol, the North Sheffield ethics committee, requesting its inclusion. This was accepted. The SF-12 was administered to a sub-set of patients at baseline who received their consultation after the ethics amendment, and to all patients at 6 month follow-up.

Despite the EQ-5D being the preferred preference based measure by NICE, the SF-12 was selected for this study as the measure is more sensitive at the higher end of the scale, and therefore to more well patients, and discussion with the lead clinician

indicated that the profile of women included in the study was relatively young and mobile women. It was therefore thought that the SF-12 would be more sensitive to changes in this specific patient population (De Smedt et al. 2014. NICE, 2013).

3.3.6 Resource use

Resource use data were also collected at two time points: the initial consultation and 6 months follow-up. A micro-costing study was undertaken on a sub-set of patients to derive costs for the initial ePAQ completion and consultation. This included all patients receiving their consultation (via telephone or face-to-face) during one of three set months throughout the trial. For these patients their consultation was timed and costs for staff time, computers, software and overheads were applied.

Six months following study recruitment, patients were posted a questionnaire asking for resources used since their consultation. Resource use data collected included personal expenditure due to bladder, bowel or vaginal problems; time off work; time away from usual activities; visits to the general practitioner (GP), nurse or other health professionals; inpatient and outpatient visits. The questionnaire also requested details of prescribed medication, however, most patients included this within personal expenditure; due to the inconsistent reporting of prescription data these costs were excluded from the economic analysis.

3.3.7 Unit costs

UK unit costs were applied to resource use estimates for patients in each study arm. All costs were inflated to 2011 United Kingdom (UK) prices using the pay and prices index and value-added tax (VAT) was excluded in line with NICE guidance (Curtis, 2011; NICE, 2013). Staff time was costed using the unit costs of health and social care, and included overhead costs (Curtis, 2011). Costs of surgical procedures were estimated using a per elective inpatient episode cost based on HRG data (National reference costs, 2012).

For those patients completing ePAQ online at home it was assumed that they would have access to a computer, and would not purchase one purely to complete this questionnaire. Those with no computer were able to arrange a time to visit the hospital to complete the questionnaire. For those patients attending for a face-to-face consultation in clinic computer costs were calculated assuming two touch screen computers with an average of 1,000 completions each per year which needed replacing every 8 years. Annual software costs were calculated assuming an average of 1,000 completions per year. These data were provided by the lead clinic consultant. The key unit costs used in the analysis are reported in Table 2.

Table 2. Unit Costs

Resource Item	Unit	Unit Cost (£)	Source	Assumptions
General Practitioner Visit	Visit	53	Curtis (2011)	Assuming clinic consultation lasting 17.2 minutes. Including direct care staff costs.
Specialist Nurse	Visit	30.50	Curtis (2011)	Assuming Nurse Team Leader, £122 per hour of patient contact (including qualifications), duration of contact 15 minutes.
Practice Nurse	Visit	12.75	Curtis (2011)	Assuming Nurse (GP practice). £51 per hour of face-to-face contact. Including qualifications Duration of contact 15 minutes.
Consultant (surgical)	Minute	2.68	Curtis (211)	Assuming Consultant, surgical, hospital based. £161 per contract hour.
Physiotherapist	Visit	13.60	Curtis (2011)	Assuming Hospital physiotherapist £35 per hour. Duration of contact for clinic appointment 23.3 minutes

Resource Item	Unit	Unit Cost (£)	Source	Assumptions
Gynaecology OP	Outpatient visit	141.00	NHS Reference costs 2010-11	NHS trusts and PCTs combined consultation led: follow up attendance multi-professional non-admitted face to face. Service code 502.
Cost per day off work	Day	88	ONS 2011	Based on 2011 Annual Survey of hours and earnings.
Computer cost in group 2 (control)	per patient	0.25	Personal communication from STH (SR)	Assuming 2 touchscreen computers costing 1,000 per computer; average number of completions 1,000; computers replaced every 8 years. Annuity factor of 6.874.
Cost of software	per patient	2.40	Personal communication from STH (SR)	Based on annual software cost of £2,400; average number of completions 1,000 per year.

Resource Item	Unit	Unit Cost (£)	Source	Assumptions
Consultant cost in group 1 (intervention) – including OH	per patient	29.35	Micro-costing study	-
Consultant cost group 2 (control) – including OH	per patient	69.52	Micro-costing study	-
Genital prolapse or incontinence	Per elective inpatient episode	1,741	HRG MB02Z	-
Hysterectomy	Per elective inpatient episode	3,346	HRG MA02Z	-
Bladder repair/bladder surgery	Per elective inpatient episode	1,958	HRG MA04B	-

Resource Item	Unit	Unit Cost (£)	Source	Assumptions
Botox to bladder	Per elective inpatient episode	1,958	HRG MA04B	-
Lower genital tract disorders without CC	Per elective inpatient episode	1,992	HRG MB01B	-

3.3.8 Cost-effectiveness Analysis

Estimates of mean costs and QALYs for the two treatment arms were calculated over the six month follow-up period. Statistical analysis was performed using SPSS. The study was undertaken from both NHS and societal perspectives. The primary analysis reports the societal perspective, and the NHS perspective is reported as a sensitivity analysis. Costs were categorised under three headings: cost of consultation, direct 6 month costs and indirect 6 month costs. Due to a six month time horizon, total costs and QALYs were not discounted.

The SF-12 was only administered to a sub-set of patients; analysis of these results was only performed on patients who had completed the SF-12 at both baseline and 6 months. For some patients, resource use data were wholly or partly missing. Base case analysis included only patients with complete data (complete case). Alternative approaches to the handling of missing data were explored in sensitivity analysis.

The cost-effectiveness of completing ePAQ online and a telephone consultation, versus completion of ePAQ and a face-to-face consultation was determined using the incremental cost-effectiveness ratio (ICER), which is estimated by dividing cost differences by QALY differences for each treatment arm. A cost-effectiveness acceptability curve (CEAC) was plotted to represent the uncertainty surrounding the cost-effectiveness results.

3.4 Results

A total of 195 women were randomised. 98 women were allocated to the intervention group, of these 92 received the intervention. Of the 6 who did not receive the intervention, 3 declined to participate post allocation; 1 attended the routine clinic; 1 suffered a family bereavement and therefore withdrew from the study. 97 patients were allocated to the control group. Of these 95 received the intervention. Of the 2 who did not receive the intervention, 1 did not attend the clinic appointment and one received an urologist referral. Baseline characteristics of patients are shown in Table 3.

Within this chapter the primary analysis is taken from a societal perspective. Despite this being against NICE recommendations, this study was not for a NICE submission, and it was considered to be more inclusive and therefore it would provide richer information than the NHS perspective. The NHS perspective was reported in the sensitivity analysis and allowed the willingness to pay threshold (which is based on an NHS perspective) to be applied (Drummond, 2005). The NHS perspective is also taken forward within this thesis in the exploration of process utility. Further detail on this is provided in Chapter 6.

Table 4 reports resource use estimates for each patient group. The independent t-test was used to compare the mean values from the two study arms. Although the cost data were found to be non-normally distributed, which can be resolved by the

utilisation of non-parametric statistical tests, others have argued that for large enough sample sizes (defined as ≥65 by Lumley et al. 2002) the independent t-test is also valid for use with non-normally distributed data (Lumley et al. 2002). Regardless of the relative validity of these two approaches, such tests are considered to be of secondary importance to this study which is concerned with overall costs-effectiveness (as represented by the ICER and the probabilities of being cost-effective at £20,000 and £30,000 per QALY) rather than differences in the amount of resources used.

Table 3. Baseline characteristics of women in each trial arm.

Characteristic	Intervention (n = 98)			Control (n = 97)		
	n	Mean	SD	n	Mean	SD
Mean age in years	95	51.51	11.20	97	50.96	10.63
Mean number of children	95	2.17	1.16	97	2.29	1.15
Mean weight (Kg)	91	71.44	12.80	93	70.39	13.44
	n	%		n	%	
Ethnic Origin			-			-
White British n (%)	90	91.8 %		95	97.9 %	
White Irish	1	1.0 %		1	1.0 %	
White Other	3	3.1 %		0	0.0 %	
Mixed Race	1	1.0 %		1	1.0 %	
Not reported	3	3.1 %		0	0.0 %	
Education level n (%)						
None	15	15.3 %		17	17.5 %	
Junior school	30	30.6 %		26	26.8 %	
Senior School	23	23.5 %		27	27.8 %	
University Education	26	26.5 %		27	27.8 %	
Not reported	1	1.0 %		0	0.0%	
	4	4.1%			0.0%	

Characteristic	Intervention (n = 98)			Control (n = 97)		
Marital Status						
Married	65	66.3%		71	73.2 %	
Cohabiting	9	9.2 %		8	8.2 %	
Single	4	4.1 %		2	2.1 %	
Widowed	3	3.1 %		5	5.2 %	
Divorced	12	12.2 %		8	8.2 %	
Separated	1	1.0 %		3	3.1%	
Civil partnership	1	1.0 %		0	0.0 %	
Not reported	3	3.1 %		0	0.0 %	

Table 4. Resource Use

	Intervention		Control			95% CI of difference		Significance
Resource	**n**	**Mean (SD)**	**n**	**Mean (SD)**	**Mean difference**	**Lower**	**Upper**	**P value**
GP Visits	73	.77 (1.112)	78	.65 (1.493)	-.113	-.312	.539	.600
Practice nurse	73	.16 (.524)	79	.16 (.517)	-.002	-.169	.164	.977
Outpatient visits	74	1.59 (1.937)	77	1.44 (1.509)	.153	-.404	.710	.588
Mean number of surgical procedures	74	.220 (0.414)	78	.180 (.386)	.037	-.092	.165	.573
Physiotherapist	74	.400 (1.030)	77	.450 (1.142)	-.056	-.406	.294	.753
Stoma nurse	74	.010 (.116)	79	.000 (0.000)	.0140	-.013	.040	.321
Incontinence nurse	74	.120 (.776)	77	.080 (.354)	.044	-.149	.236	.655
Nurse specialist gynae	76	.030 (.161)	78	.000 (.000)	.026	-.011	.063	.159
Consultant (f2f)	75	.170 (.601)	77	.260 (.616)	-.086	-.281	.109	.383
Personal expenditure in 6 month follow-up period (£)	74	27.944 (53.207)	73	15.503 (27.666)	12.441	-1.395	26.294	.078

	Intervention		Control			95% CI of difference		Significance
Resource	**n**	**Mean (SD)**	**n**	**Mean (SD)**	**Mean difference**	**Lower**	**Upper**	**P value**
Time off work (days)	70	3.09 0 (12.473)	78	5.210 (20.763)	-2.119	-7.763	3.525	.459
Time away from normal activities	74	2.410 (8.039)	79	1.290 (6.752)	1.114	-1.252	3.480	.354

Estimated mean costs per patient for each treatment arm are reported in Table 5. Consultation costs for those patients completing ePAQ online and receiving a telephone consultation (intervention group) were less than half the costs of the group receiving standard care (control) (£31.75 versus £72.17). The key driver behind this was the duration of the consultation with the doctor and associated labour costs. The mean duration of the telephone consultation was 10.94 minutes, compared to a mean duration of 25.9 minutes for patients attending a face-to-face consultation.

Direct costs incurred during the 6 month follow-up period differed between treatment arms, with patients in the intervention group incurring greater direct costs in comparison to the control group. This was driven primarily by the difference in costs associated with gynaecology outpatient attendances between arms, with those in the intervention group incurring costs of £62.67 greater than those in the control.

Personal expenditure during the 6 month follow-up was also higher in the intervention group. However, less costs associated with loss of productivity for the intervention group resulted in lower total indirect costs per patient. The mean total cost per patient was estimated to be £1,139.86 for patients receiving the intervention and £1,101.82 for the control group. This resulted in a mean differential cost of the intervention versus control of £38.04.

Table 5. Mean Costs per Patient for Intervention and Control Groups

Resource	**Intervention (£)SD N = 27**	**Control (£) SD N = 30**
Cost of consultations		
Consultation cost	29.35	69.52
Cost of software	2.40	2.40
Cost of computer	0.00	0.25
Total consultation costs per patient	**31.75**	**72.17**
Direct costs during 6 month follow-up		
GP Visits	41.22 (49.49)	35.33 (65.78)
Practice nurse	0.94 (3.40)	2.13 (5.88)
Outpatient visits	250.67 (316.09)	188.00 (246.547)
Cost of surgical procedures	330.44 (707.375)	285.63 (784.36)
Other professionals		
Physiotherapist	5.04 (12.59)	4.99 (15.76)
Specialist nurse (including stoma nurse, incontinence nurse, gynaecology)	4.52 (18.35)	2.03 (11.14)
Consultant (f2f)	77.94 (24.44)	14.29 (37.06)
Direct costs: 6 months	**640.77 (844.40)**	**532.41 (867.09)**
Indirect costs during 6 month follow-up		
Personal expenditure in 6 month follow-up period (£)	24.07 (31.05)	16.17 (20.97)
Loss of productivity	443.26 (1573.15)	481.07 (1475.01)
Indirect costs: 6 months	**467.33 (1569.42)**	**497.24 (1479.79)**
Total costs per patient (Societal perspective)	**1,139.86 (2182.24)**	**1101.82 (2172.44)**

Table 6 reports SF-6D results for all patients who completed the SF-12 at both baseline and 6 months. Within the intervention group patients mean utility estimates reduced slightly from baseline to 6 months. Equivalent estimates for the control resulted in a slight increase in quality of life over the 6 month period. These estimates resulted in a QALY loss for patients in the intervention group, compared to QALY gains for patients receiving standard care. No differences were found to be statistically significant.

Table 6. Mean utility per patient for the intervention and control groups.

Item	Intervention (N=27)	Control (N = 30)		95% CI of difference		
	Mean (SD)	Mean (SD)	Difference (SE)	lower	upper	Significance
SF-6D baseline	.64 (0.90)	.62 (.081)	.026 (.024)	-.022	.0744	.287
SF-6D 6 months	.63 (.082)	.62 (.091)	.00698 (.023)	-.039	.0531	.763
Change in SF-6D	-.0152 (.073)	.0038(.094)	-.01899 (.02245)	-.0640	.0260	.401
QALYs gained	-.0076 (.037)	.0019 (.047)	-.0095 (.1122)	-.3199	.0130	.401

As both the costs and QALY loss for the intervention were slightly higher than the control, the cost and utility point estimates from a societal perspective indicate that the control treatment dominates.

3.4.1 Deterministic Sensitivity Analysis

NICE (2013) stipulates that cost-effectiveness analyses are undertaken from an NHS perspective. As this EEACT was not undertaken for a NICE appraisal, the primary analysis was performed from a societal perspective, and analysis from an NHS perspective was analysis was undertaken as a sensitivity analysis. This included only the initial consultation costs in addition to the direct 6 month costs. This again resulted in slightly higher costs for the intervention group (£672.52), when compared to the control group (£604.58), and a negative ICER. Therefore, the control remained dominant over the intervention.

Additional sensitivity analysis could have been undertaken, and is routinely undertaken within economic EEACTs (Drummond, 2005). For example, one-way sensitivity analysis could have been performed to understand the relationship between single parameters on the outputs, and multivariate analysis could have been used to understand the impact of altering more than one parameter. These analyses were not included within the original trial protocol (devised by STH and ScHARR), which gained ethics approval, and therefore were not performed by VB.

3.4.2 Probabilistic Sensitivity Analysis

A further analysis was undertaken from an NHS perspective and which accounted for missing data. Initially, V.B. performed an analysis which used a non-parametric bootstrap of costs and QALYs. However, the analysis was re-run using updated methods, with code written by Abu Alshreef, from ScHARR.

The analysis was performed in line with the ISPOR Guidelines for undertaking EEACTs in which no specific approach is recommended, however, transparency is advocated. They suggest that where there are missing data, the nature of the missing data should be determined, and an approach for dealing with the missing data should be defined (Ramsey et al. 2015). Except where there are only small levels of missing data <5% for one parameter, Ramsey et al. (2015) recommend against excluding all cases with missing data, and advocate the use of multiple imputation methods.

A descriptive analysis of NHS costs and QALYs was undertaken to identify the amount of missing data. 146 observations were available for NHS costs and 59 observations available for QALYs. The much higher rate of missing SF-12 data was due to a clerical error that lead to the SF-12 not being administered in the first few months of recruitment. Missing data might result in misleading cost-effectiveness estimates, whilst complete case analysis is undesirable as it might result in reducing the number of observations included in the analysis, and hence, affect the power of the study. Examination of the pattern of missing data and knowledge of the reason for the missing SF12 data suggested that these data were missing completely at random. As such, imputation was considered appropriate and among the available methods, multiple imputation was considered to be the best available method (Faria et al. 2014).

The MI command within Stata® was used, with the number of imputations set equal to the rate of missing data within the least complete variable (Faria et al. 2014; Gomes et al. 2012); which led to 70 imputations being used. Costs and change in QALYs from

baseline were jointly imputed using patient age, and parity and treatment group as explanatory imputation variables. The imputation was performed by treatment group for all imputed variables except baseline utility. Imputation of baseline utility was performed across all observations rather than specified by treatment group. The MI process generates the imputed data by using the Multiple Imputation Chained Equation (MICE) with Predictive Mean Matching (PMM) method (Faria et al. 2014). This model generated imputed values by prediction of missing data from the posterior distribution of missing observations given the values of observed data from the imputation variables.

The SUR was then undertaken across the 70 resultant datasets in order to take account of the uncertainty in the imputed values, but in a way that recognises that the datasets do not represent independent observations. Likewise, the cost-effectiveness confidence ellipses and cost-effectiveness acceptability curves (CEAC) were generated (parametrically) using means, correlations and standard errors generated from the output of the SUR estimated across the 70 datasets (Hoch et al. 2002). Five parameters were used in this analysis: the difference in mean total cost, standard error of the difference in mean total cost, difference in mean QALYs, standard error of the difference in mean QALYs and the covariance between cost and QALYs. These parameters were used for producing the CEAC and confidence ellipses.

QALYs were calculated as change from baseline. Whilst some researchers argue for the statistical adjustment of QALYs using baseline utility within a regression, this is

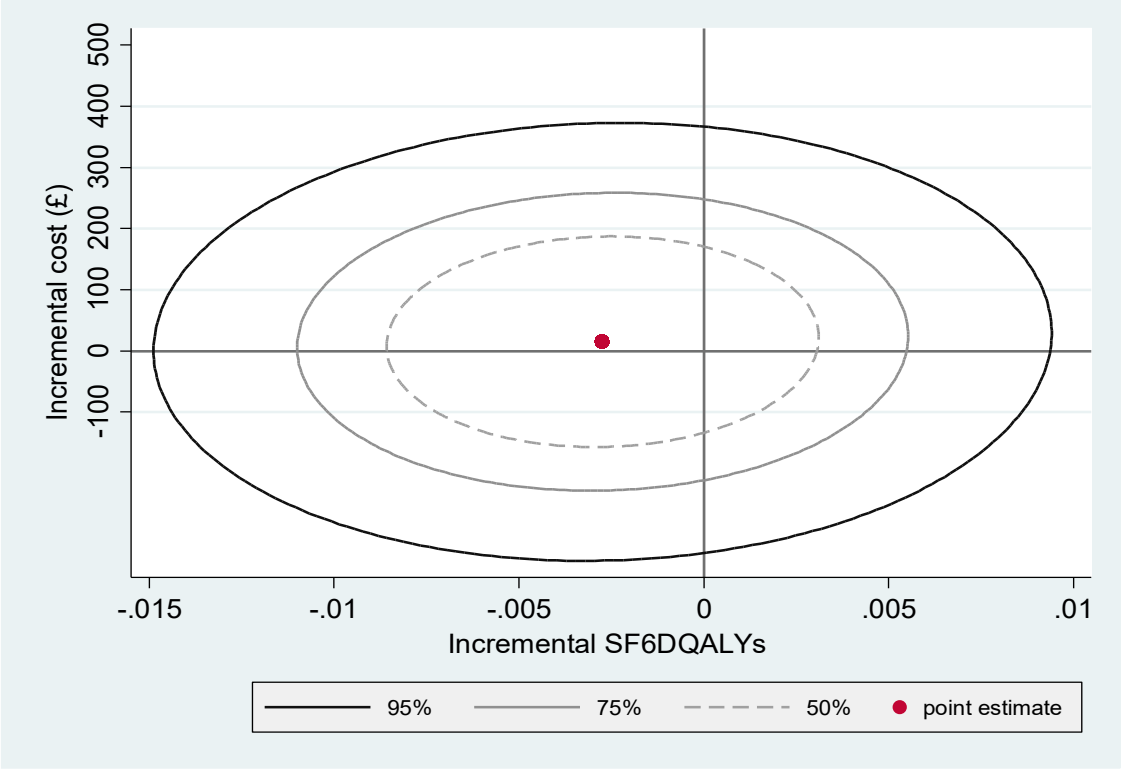
problematic in the context of this study. This is because the comparison between the SF-6D and SF-7D analysis (a bolt-on to the SF-6D that includes a process domain and will be explained further in Chapter 5) will be a combination of the statistical adjustments for the two regressions, plus the differences in the calculated utilities. Using change from baseline removes the effect of a differential statistical adjustment, thereby giving a clearer picture of the effect of using the two different utility calculations.

Outputs from the SUR can be seen in Table 7. The table shows the mean values (cost and utility) for the control group, and the mean incremental value for the intervention. Therefore the analysis shows additional costs (£15.10) and lower QALYs (0.003) in the intervention group, with p-values of 0.918 and 0.582 respectively. Figure 11 shows that these point estimates sit in the north-western quadrant, however, the 95% confidence ellipse spreads to all four quadrants. As illustrated by Figure 12, there is approximately a 35% chance that the intervention is cost-effective at a threshold of £20,000 per QALY.

Table 7. Cost and SF-6D

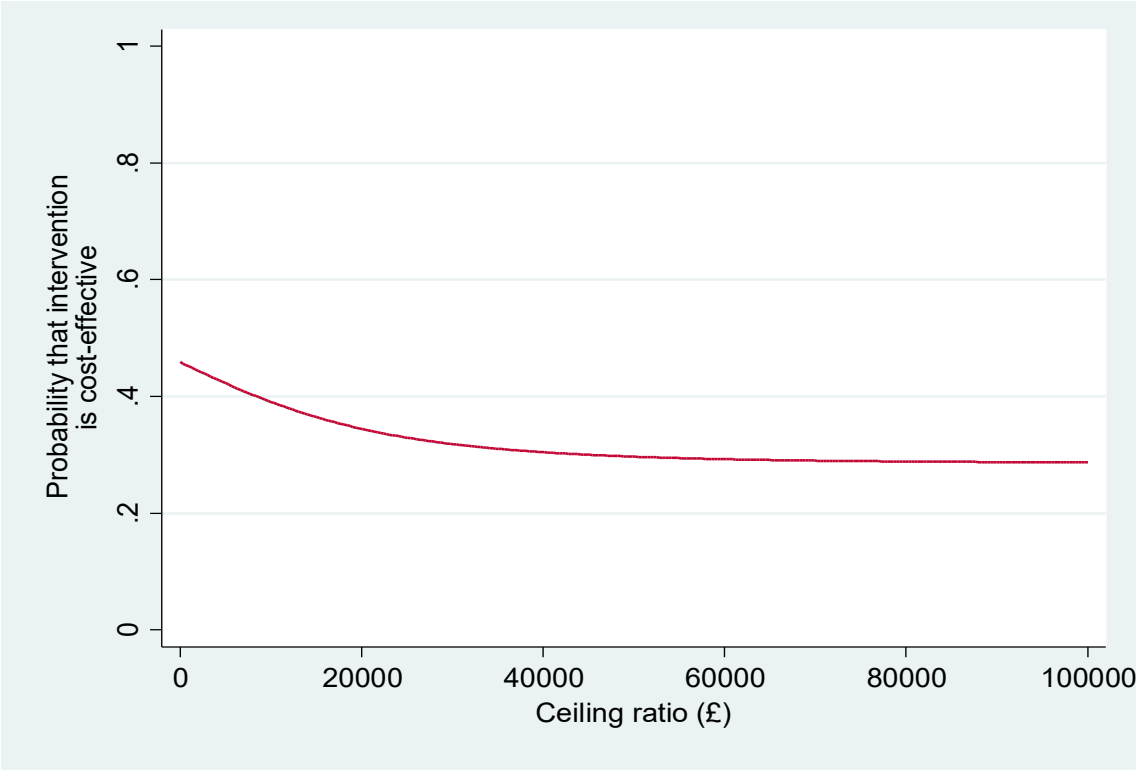
	Group	Co-efficient	Std.err	t	P>	95% CI	
NHS Costs	Intervention[1]	15.10	146.30	0.10	0.92	-271.87	302.08
	Control[2]	701.40	101.42	6.92	0.00	502.50	900.30
SF-6D[3]	Intervention[1]	-.00273	.00496	-0.55	0.58	-.01251	.00705
	Control[2]	.00259	.00354	0.73	0.47	-.00440	.00957

Figure 11. Cost Effectiveness Plane.



[1] Mean incremental value (compared to control)
[2] Mean value for the control group
[3] Change from baseline.

Figure 12. Cost-effectiveness acceptability curve



3.4.3 Summary of clinical results

Analysis of the clinical outcomes was performed by Dr Georgina Jones, and Richard Jaques from ScHARR at The University of Sheffield. The primary outcome measure was the patient experience questionnaire (PEQ).

Analysis of the PEQ found no differences in short term outcomes between the intervention group and the group receiving standard care. However, there was evidence to suggest a statistically significant difference between the control and intervention groups for the other three dimensions of the PEQ (communications, emotions and barriers).

3.5 Discussion

This economic evaluation aimed to estimate both the direct and indirect costs of the introduction of a "virtual clinic" in urogynaecology. The primary analysis was reported from a Societal perspective. Although cost-effectiveness analyses for NICE appraisals require an NHS perspective (NICE, 2013), this study was not performed as part of a NICE appraisal, and therefore it was deemed more relevant and informative to include both direct and indirect costs. However, as the cost-effectiveness threshold is based on an NHS perspective, this perspective was used in the sensitivity analysis for the handling of missing data.

The study found that the costs associated with the initial consultation were considerably less in the intervention group (£34.73) compared to the control group (£77.82). This difference was due to the duration of the consultation with the doctor and associated labour costs. The mean duration of the telephone consultation was 10.94 minutes, compared to a mean duration of 25.9 minutes for patients attending a face-to-face consultation.

Direct costs incurred during the 6 month follow-up period differed slightly between treatment arms, with patients in the intervention group incurring slightly higher costs (£640.77 for the intervention group, and £543.41 for the control). Patients in the control group had slightly higher total indirect costs.

Total costs indicated that at 6 months post consultation, the combination of completing ePAQ online and receiving a telephone consultation resulted in a cost saving per patient of £68.63 when compared to standard care. However, it also resulted in a slight reduction in patient's quality of life. The intervention therefore lies within the North West quadrant of the cost-effectiveness plane indicating that it is more expensive and less effective, and therefore dominated by standard care. From a societal perspective therefore, the introduction of the service into the NHS would therefore be at the discretion of decision makers (Gray et al. 2006). When considering an NHS perspective, the costs for the intervention remain slightly higher than the control (£672.52 and £604.58 respectively), and as shown by the point estimate in Figure 11, the intervention remains in the North West quadrant, and dominated by standard care.

The SUR analysis accounted for both the handling of missing data, and for correlation between costs and QALYs. This again resulted in higher costs (by £15.10) and lower QALYs (by 0.003) in the intervention group; however, Figure 11 provides an indication of the levels of uncertainty in estimates as the 95% confidence ellipse spreads to all 4 quadrants, with only approximately 35% probability that the intervention is cost-effective at a £20,000 per QALY threshold (Figure 12).

There are several points to consider when reviewing the study results. Comparison of the consultation costs alone clearly indicated cost savings associated with the intervention. However, the inclusion of wider NHS and societal costs lead to higher

mean costs for the same group, although this did not meet statistical significance. It could be that the inclusion of these costs increased the amount of variability between cost comparators in a study that was not statistically powered to detect differences in cost endpoints. A more appropriate approach may be to consider only the direct costs associated with the consultations.

The study also resulted in slightly lower health related quality of life for the intervention group. Whilst it is standard practice to use a generic preference based measure such as the SF-6D to obtain the utility estimates required for input into the QALY, it is unlikely that it would have captured any differences between patients' preferences for these specific interventions, given that the SF-6D contains only information on health outcomes, and does not encompass any aspects of preference for client satisfaction or assessment of the service delivery. It could therefore be argued that it is an inappropriate choice of outcome, which is unlikely to capture differences in patients' satisfaction with the service and is potentially the reason for such small changes in QALYs between the two treatment arms. Unless a measurement of patient satisfaction can be built into the QALY then the trial results may be providing an inaccurate estimate of cost-effectiveness for this comparison of the processes of receiving health care.

3.5.1 Strengths and Limitations.

The findings from the ePAQ EEACT will now be discussed in relation to the threats to validity which were identified in Chapter 2.

This economic evaluation was undertaken alongside a randomised controlled trial. The study therefore benefited from the randomisation of patients between groups and the reduction in the potential for bias this brings (Ramsey et al. 2015). The intervention was compared to standard care at STH, therefore allowing an ICER to be estimated to show the cost-effectiveness of introducing this new treatment approach when compared to routine practice. However, the use of ePAQ online or in advance of clinic appointments is not routine throughout the UK, therefore this comparator only produces estimates which are informative locally in resource allocation decisions. The introduction of a comparator where no patients are given ePAQ to complete would mean removing it from those receiving standard care and was thought to be unethical, and therefore not included in the trial design. Given that cost differences were identified between study arms, and the introduction of the intervention may therefore have an impact on resource allocation, it may be appropriate to run the study in a different location where standard care does not include the completion of ePAQ.

The lack of generalisability of the trial results is likely to be further compounded by the trial setting. STH is a teaching hospital, where the costs incurred are unlikely to reflect those of a community or district hospital. For example, patients attending clinics at teaching hospitals are more likely to see a consultant than those attending district hospitals. The effectiveness of ePAQ as a tool for triage also depends on the patient pathways and referral systems already in place in the hospital. These systems may be less developed within the community hospital setting, and may differ between

teaching hospitals. The results from the RCT may not therefore be wholly generalisable to the UK health service.

Despite the potential for the results to be non-reflective of standard care throughout the country, the study design did include broad patient inclusion criteria. All women aged 18 or older referred to the urogynaecology clinic, who were able to understand and speak English, were eligible for inclusion. Although excluding non-English speaking patients makes the sample less likely to be a true representation of the national population, it does not impose major restrictions on entry to the study (Rittenhouse 1996).

The study also has a potential bias due to the impact that the intervention has on the speed of referral and subsequent treatments; a telephone consultation may lead to quicker referral on to other services, which could lead to higher costs within the 6 months follow-up which may even out over the long term. The study may therefore be limited by the duration of the follow-up, and benefit from a duration long enough to capture these differences (O'Brien et al. 1996).

This economic evaluation aimed to estimate both the direct and indirect costs of the new approach to service delivery. Whilst the costs associated with time off work and time spent away from normal activities were accounted for over the 6 month follow-up duration, personal transportation costs for attending the clinic consultation (both monetary and time) were not considered. Patients who completed ePAQ in clinic and received a face-to-face consultation are likely to have incurred greater costs for the

time and expense of travelling to the hospital. This omission is likely to bias results against the intervention. Data on medication taken over the 6 month follow-up period were collected. However, in reviewing the completed and returned questionnaires, some patients had included this information in the “out of pocket expense” question, and others in the “details of prescribed medications” question. Due to this inconsistency in the reporting of resource use a decision was therefore taken to exclude medication costs from the analysis, due to potential inaccuracies in the interpretation of responses. Travel expenses were not included as the original investigator (SD) considered that these costs, which are only relevant from the societal perspective would be negligible in comparison to the value of patient time. There is also uncertainty about the value placed on transport as patients tend to do other things whilst coming to the city centre, e.g. shopping, so it is difficult to identify a 'pure' cost of transport. Given these issues and a desire to reduce patient burden, it was decided not to ask about transport costs.

No data were collected on the time patients took to complete ePAQ questions and the questionnaire as a whole, although online completion facilitates its collection. As patients in both trial arms were completing the questionnaire online, and in advance of either a telephone or face-to-face consultation no differences in time to complete ePAQ were expected.

The SF-12 was only completed at 2 time-points: immediately before the consultation and at 6 months following the consultation. In order to gain a more specific measure of

the utility associated with the consultation alone, it could also have been appropriate to administer it immediately following the consultation. However, the project team had concerns that this would introduce questionnaire completion fatigue and there was also a concern of the risk that the SF-12 may have contaminated the QQ-10 responses: The QQ-10 asks respondents about the questionnaire they have just completed, and was aimed at gaining their views on ePAQ completion. The addition of more questionnaires may therefore have caused confusion, and take the focus off ePAQ. The utility associated with the consultation alone may therefore have been more accurately estimated if questionnaires were collected at 3 time-points, however, this may have been at the detriment of the QQ-10, In addition to this, retrospectively, it is clear that there would have been an issue with the number of patients with complete SF-12 data, as when collecting at 2 time-points there were only 27 in the control.

When considering this study in relation to the threats to validity identified in Chapter 2 therefore, it is evident that only issues with the comparison therapy and length of follow-up were identified. However, in addition to these it is also evident that the measure of the benefits using standard cost-effectiveness methods was not designed to capture all of the relevant effects of the technology.

These cost-effectiveness results follow the same trend as shown in other studies. For example, Pinnock et al. (2007) compared the cost-effectiveness of face-to-face nurse lead consultations, with telephone consultations in patients with asthma. They found

that the telephone consultations were significantly shorter (11.19 minutes for telephone consultations, versus 21.87 minutes for those attending face-to-face). They also found total costs per patient over the follow-up period (3 months) to be similar between groups. Their study also therefore indicated that the introduction of telephone consultations could lead to cost savings in the NHS, although the differences were not as defined with the inclusion of wider NHS costs. Their study was undertaken from an NHS perspective, and therefore societal costs were excluded. Their inclusion may have resulted in even more favourable estimates of cost-effectiveness for the intervention group, due to the saved time and expense of travelling to hospital appointments. Pinnock et al. (2007) also included outcome measures related to patients' quality of life (Mini-asthma quality of life questionnaire: Mini-AQLQ), and enablement (Modified Patient Enablement Instrument: mPEI). Just as the ePAQ EEACT collected data on quality of life (SF-12) and patient satisfaction (PEQ, CSQ-8). Both studies therefore provided a comprehensive overview of the impact of the alternative consultations on patients.

The research to date is not however conclusive. Beaver et al. (2009a) performed an EEACT which compared the costs associated with a face-to-face follow-up consultation to telephone follow-up after treatment for breast cancer. They found that patients who received telephone follow-up had a longer duration of consultations which lead to higher costs. Their cost-minimisation study also included patient and carer travel and productivity costs and found that these were significantly higher for patients who had hospital-based follow-up. Although Beaver et al. (2009) report only the methods and

results for the cost-minimisation study, outcome measures including the State-Trait Anxiety Inventory (STAI), and patient satisfaction measures (exact measure not reported) were also included, however reported in a separate clinical publication (Beaver et al. 2009b). Both Pinnock et al. (2007), and Beaver et al. (2009a) therefore, also consider patient satisfaction, and the impact of the consultations on patients quality of life, however, neither also include the QALY as an outcome.

3.5.2 Recommendations for future research

This study has identified points to consider for future work. These include studies performed in a setting where the use of ePAQ is not standard care; and studies with longer-term follow-up which would provide greater insight into the impact of ePAQ online and telephone consultations on time to treatment and the associated impact on longer term costs.

Subsequent studies should also consider more fully the short term costs incurred by patients, including not only the time away from work, but also the costs and time associated with attending hospital appointments.

It is also important that, when designing and undertaking cost-utility studies all of the relevant outcomes of an intervention are considered, whether they be health outcomes or outcomes associated with the processes of receiving care.

3.6 Conclusion

This cost-effectiveness analysis indicated that the introduction of ePAQ online in advance of a telephone consultation was unlikely to be cost-effective at the £20,000 threshold. However, the inclusion of costs which were not directly related to the consultation itself (e.g. surgical costs within the 6 month follow-up) introduced uncertainty. It may therefore be sensible to ignore these wider costs, and concentrate solely on the cost of the consultation, following the approach of Beaver et al. (2009). In order to obtain a more detailed insight of the cost implications for the intervention types, the societal costs to the patient in terms of travel time and costs should also be considered.

When considering the measurement of benefits, given that we expected no differences in the health outcomes between patient groups, the standard utility measure used (SF-6D), which captures differences in patient's preferences for health outcomes, was unlikely to capture any differences associated with the mode of consultation. Both the SF-6D and wider NHS costs were found to be irrelevant in this study to evaluate the cost effectiveness of 2 differing processes of care. If the QALY is to be used as a generic measure of benefit across cost-utility studies, including which impact on both health and process outcomes, further work should be undertaken to expand on its current form to ensure all of these relevant outcomes are captured, resulting in more accurate estimates of cost-effectiveness.

3.7 References

- Beaver, K. Hollingworth, W. McDonald, R. Dunn, G. Tysver-Robinson, D. Thomson, L. Hindley, AC. Susnerwala, SS. Luker, K. (2009a) Economic evaluation of a randomised clinical trial of hospital versus telephone follow-up after treatment for breast cancer. *British Journal of surgery*, 96(12), 1406-1415.

- Beaver, K. Tysver-Robinson, D. Campbell, M. Twomey, M. Williamson, S. Hindley, A et al. (2009b) Comparing hospital and telephone follow-up after treatment for breast cancer: randomised equivalence trial. *BMJ* 338: a3147

- Brazier, JE. and Roberts J. (2004) The Estimation of a Preference-Based Measure of Health from the SF-12. *Medical Care,* 42(9), 851-859.

- Curtis L. (2011) Personal Social Services Research Unit: Unit costs of Health and Social Care 2011. Available from: http://www.pssru.ac.uk/archive/pdf/uc/uc2011/uc2011.pdf [Accessed 22nd February 2013].

- De Smedt, D. Clays, E. Annemans, L. De Bacquer, D. (2014) EQ-5D versus SF-12 in coronary patients: are they interchangeable? *Value Health*, 17(1),84-9.

- Drummond, MF. Schulpher, MJ. Torrance, GW. O'Brien, B. Stoddart, GJ. (2005) Methods for the Economic Evaluation of Health Care Programmes. 3rd Edition, Oxford, Oxford Press.

- Faria, R. Gomes, M. Epstein, D. White, IR. (2014) A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics*, 32(12),1157-70.

- Field, A. (2009) *Discovering Statistics Using SPSS*. 2nd Ed. London, Sage Publications Ltd.

- Gomes, M. Grieve, R. Nixon, R. Ng, ES. Carpenter, J. Thompson, SG. (2012) Methods for covariate adjustment in cost-effectiveness analysis that use cluster randomised trials. *Health Econ.* 21(9),1101-18.

- Gray, A. (2006) Cost-effectiveness analyses alongside randomized clinical trials. *Clinical Trials*, 3,538-542.

- Hoch, JS. AH, Briggs. and AR, Willan. (2002) Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Econ*, 11, 415–430.

- Jones, GL. Radley, SC. Lumb, J. Jha, S. (2008) Electronic pelvic floor symptoms assessment: tests of data quality of ePAQ-PF. *Int Urogynecol J Pelvic Floor Dysfunct,* 19(10),1337-47.

- Larsen, DL. Attkisson, CC. Hargreaves, WA. Nguyen, TD. (1979) Assessment of client/patient satisfaction: Development of a general scale, *Evaluation and Program Planning*, 2, 197-207.

- Lumley, T. Diehr, P. Emerson, S. Chen, L. (2002) The Importance of the Normality Assumption in Large Public Health Data Sets. *Annu Rev Public Health* 23,151-69.

- Manca, A. Hawkins, N. Sculpher, MJ. (2005) Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ,* 14(5),487-96.

- National Schedule of Reference Costs (2011-12) NHS Trusts and PCTs Combined. Department of Health. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-financial-year-2011-to-2012 [Accessed October 2011].

- NICE (2013) Guide to the methods of technology appraisal. Available from: https://www.nice.org.uk/article/pmg9. [Accessed August 2014].

- O'Brien, B. (1996) Economic Evaluation of Pharmaceuticals. Frankenstein's monster or Vampire of trials? *Medical Care*. 34(12)DS99-DS108 Supplement.

- ONS (2011) Office of National Statistics. 2011 Annual Survey of Hours and Earnings. Available from: http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings/ashe-results-2011/ashe-statistical-bulletin-2011.html. [Accessed December 2011].

- Pinnock, H. Adlem, L. Gaskin, S. Harris, J. Snellgrove, C. (2007) Accessibility, clinical effectiveness, and practice costs of providing a telephone option for routine asthma reviews: phase IV controlled implementation study. *Br J Gen Pract,* 1,57(542),714-722.

- Ramsey, SD. Willke, RJ. Glick, H. Reed, SD. Augustovski, F. Jonsson, B. et al. (2015) Cost-effectiveness analysis alongside clinical trials II-An ISPOR Good Research Practices Task Force report. *Value Health,* 18(2),161-72.

- Rittenhouse, B. (1996) Uses of models in economic evaluations of medicines and other health technologies. *Office of Health Economics*, BSC Print, London.

- Steine, S. Finset, A. & Laerum, E. (2001). A new, brief questionnaire (PEQ) developed in primary care for measuring patients' experience of health interaction, emotion and consultation outcome. *Family Practice,* 18(4), 410-417.

Chapter 4: Process utility and a Literature review.

This chapter explores process utility in more depth, and is routed in the literature. It begins with a discussion on the concepts of process utility, and then reports a systematic literature review into studies reporting an empirical measure of process utility. The different approaches used, and utility estimates are reviewed, and the findings are used to inform subsequent chapters in the development of a process utility valuation study.

4.1 Introduction

As detailed in Chapter 2, which provided an overview of the economic methodology underlying this thesis, within health economics research it is standard practice to seek to maximise health gain within a given budget constraint. Health economics research in England has evolved to a position where this is achieved through the maximisation of the quality-adjusted life-year (Oliver, 2012). As the QALY is the dominant approach to economic evaluation in England, this has remained the focus throughout this thesis. Economic evaluation, as routinely applied to health care, assumes that individuals derive utility only from the consequences or outcomes of actions or processes, and not from the actions or processes themselves (Brazier et al. 2007). This assumption of consequentialism suggests that consumers of health care gain no benefit from its consumption, but that its value is determined solely by the health outcomes it generates (Mcguire et al. 1988). The health consequences experienced by the patient

are therefore considered to be the most important element of the patient's health care utility function (Mooney 1998). However, it has been suggested that health economists ignore the processes that patients have to go through in the course of their treatment, and that these processes should also be incorporated into the utility function (Mooney 1998; Donaldson and Birch, 2004). Included in these are the "process" of treatment and additional consumption benefits such as re-assurance or information (Donaldson and Shackley 1997). Their inclusion would take into account that the process of receiving care impacts on patients' utility, as well as the health gains they achieve (Gerard and Mooney 1993).

Several authors have published opinions on the definition of process utility. McAlister et al. (1994) suggested that the utility function should include three dimensions: the process of care; the health outcome; the structure of the providers of care (for example, including the resources, physical and organisational settings). Donaldson and Shackley (1997) suggested that including only the benefit that arises from health care in the consumers utility function leads to "narrow consequentialism". Mooney (1998, p.101) defines "narrow consequentialism" as "the monopolization of the utility or welfare function…by health and health alone." This has led to researchers ignoring other consequence and process issues. Mooney (1998 p.102) justifies the existence of process utility through the notion that processes are not always "utility neutral" but that patients can experience negative utility when receiving the healthcare needed to gain the health outcomes). Donaldson and Shackley (1997) reiterate this opinion by suggesting that historically it has been assumed that health care has no positive value

in use, however, this ignores the possibility that health care has a negative value in use. When related to practice it seems erroneous to assume consumers would have no preference for less-invasive laparoscopic surgery versus open surgery if health outcomes were equal. McAlister (1994) further iterates this by suggesting that if non-health outcomes are included in the utility function (such as provision of information and relief from anxiety), patients' utility may increase while their health status remains unchanged. Brouwer et al. (2005) relate the concept of process utility to caring for the chronically ill, where treatment is not primarily aimed at health gains. To overcome this limitation to the utility function Mooney (1998, p.102) proposes a "pluralistic benefit function" which incorporates all relevant outcomes and processes.

Mooney (1998, p.99) suggests that it is not unexpected that the primary focus of health service evaluation is health outcomes, however, he introduces the question "are there no other effects that health services have?" , which poses the question "Are the consequences of health care too narrowly defined?". Donaldson & Shackley (1997) suggest that it is not for health care analysts to place restrictions on what enters people's utility functions; by doing so economists may be reporting cost-utility analyses which present inaccurate conclusions and which are then used by policy and decision makers to determine inefficient resource allocation. Relating this to the online use of ePAQ prior to clinic appointments, cost-utility analyses may be reporting less favourable estimates of cost-effectiveness by failing to capture differences in patients utility caused by the process of receiving care. Tavakoli et al. (2000, p. 116) relate the concept of process utility specifically to decision analysis models, which compare the

"expected pay-off" of one treatment route over another . They suggest that the focus of decision analysis models on net expected values ignores components of patients care such as keeping their treatment options open.

As suggested by McAlister (1994, p.21) therefore, "the challenge is to develop approaches to the conceptualization and measurement of quality that are both clinically relevant and meaningful for policy-makers, patients and managers."

4.1.1 Defining Process Utility

As stated by Donaldson & Shackley (1997, p.699), the initial issue to overcome in a study measuring "process utility" is the definition of process utility itself. They suggest that the "process" of treatment may affect utility, but that there are additional difficulties in classifying attributes of health care into "process" and "outcome". Within their study (1997, p.700) they define health as "the effects of health care on life years and on functional aspects of health related quality of life". They then include additional attributes of health care which may impact on utility such as dignity and re-assurance, and define these as "health attributes". Finally they include non-health aspects of care such as receiving surgery or drugs, and define this as "process utility". Swan and Sainfort (2003, p.267) contradict this definition by suggesting that "testing and other short-term events can also be viewed in terms of "process utility" which attaches utility or value to aspects that are proximal to eventual outcomes, such as information, reassurance, and morbidity" . These two examples also highlight anomalies in authors' categorisation of the components of the utility function, as Donaldson and Shackley

(1997) categorise “reassurance” as a “health attribute” but Swan and Sainfort (2003) categorise it as an aspect of “process utility”. This argument is developed further by Gonzalez Block et al. (2001) who performed a study to test the hypothesis that population groups view health gain as only one among several benefits derived from health systems, and reported that study respondents distinguished between process benefits whose value ends once service interaction has ceased and reassurance-related benefits whose value was more enduring.

Salkeld (1998, p.106) puts forward a definition of process attributes and suggests that “if ‘health’ is narrowly defined purely in terms of final morbidity and mortality outcomes that accrue due to preventative activities, then process attributes can be defined as utility-bearing characteristics that occur up until the final outcomes are obtained.”

McAlister (1994) provides further discussion on measurement of quality on health care. She introduces 3 approaches to the assessment of quality put forward by Donabedian (Donabedian, 1988, cited in McAlister, 1994) which include assessment of structure, process and outcome. However, her examples of “process” include the ethics and values of society as opposed to the non-health aspects of care such as drugs and receiving surgery included under “process” by Donaldson and Shackley (1997).

Despite the lack of consensus in determining what components of health care should be included in the “pluralistic benefit function”, the reviewed literature identifies the 4 main elements illustrated in Figure 13.

Figure 13. A Venn Diagram illustrating a "pluralistic benefit function".

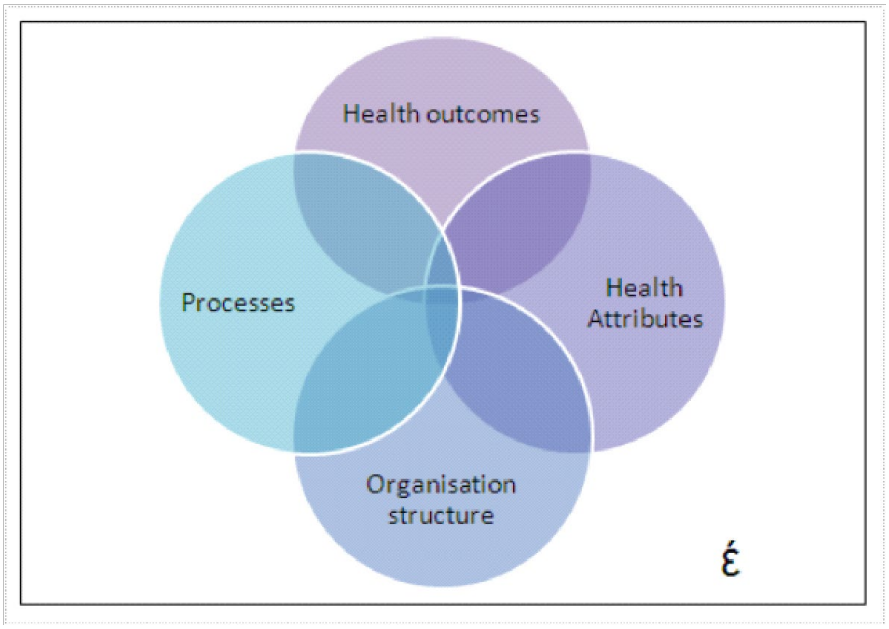


Table 8 provides a summary of the papers that have attempted to define what aspects of utility should be included in the pluralistic benefit function. For the purpose of this study the classification provided by Donaldson & Shackley (1997) was adopted. At the time, I felt that their categorisation of non-health aspects of care such as drugs and receiving surgery as "process" was in line with my own interpretation and categorisation of process utility. They included:

1. Health outcomes: e.g. effects of health care on life years and functional health;
2. Processes: e.g. non-health aspects of care such as drugs and receiving surgery;
3. Health attributes: e.g. dignity and reassurance.

They excluded the "organisational structures" proposed by McAlister (1994) but reported a more inclusive utility function than Swan & Stainford (2003) who only included testing and short term events, information and reassurance.

Table 8. Defining the Utility Function.

Author (year of publication)	Processes	Health Outcomes	Health Attributes	Organisational structure
Donaldson & Shackley 1997	✓ e.g. non-health aspects of care, such as receiving surgery or drugs.	✓ e.g. effects of health care on life years and functional HRQOL	✓ e.g. dignity, reassurance	**X**
McAlister 1994[1]	✓ e.g. the ethics and values of society	✓ e.g. the change in a patient's health	**X**	✓ e.g. physical and organisational setting
Swan & Stainford (2003)[2]	✓ e.g. testing and short term events; information, reassurance	**X**	**X**	**X**

1 Based on a study by Donabedian (1998)

2 Swan & Sainfort (2003) provide discussion on process utility alone, no additional components of the utility function are reported.

The concept of process utility has been introduced and defined, and this thesis now moves on to explore published work which has attempted to identify and measure process utility within health care.

The thesis focuses on the QALY, and methods to adapt or change the way in which the QALY gain is estimated, as opposed to using alternative methods such as willingness to pay. This decision was made because the dominant approach to economic analysis in England is the QALY. Although it has limitations (not discussed in detail in this thesis), it is used, and understood by clinicians and commissioners, and is the approach recommended by NICE. It was therefore felt that advancing our knowledge in relation to the QALY would provide the most valuable information to the field.

4.2 A literature review of process utility studies.

4.2.1 Aims and Objectives

- To identify studies which produce an empirical measure of process utility.
- To identify studies which produce process utility estimates which are compatible with the QALY.
- To explore the approaches taken and methodologies used to inform a process utility valuation study linked to the ePAQ EEACT.

4.2.2 Methods

Search Strategy

The search strategy was based on 4 combinations of free text terms; these are reported in Table 9. Search 1 comprised only of the term “process utility”. However, it was important to identify not only those studies specifically reporting “process utility” but also studies which were designed and conducted to provide a measure of the “process” of health care, but which did not use the “process utility” description. Therefore a further 3 searches were performed in order to capture more studies reporting a measure of process utility. Initial searches were performed in Pubmed in April 2013, and updated in August 2015 using both Pubmed and Web of Science.

Table 9. Search Strategy.

Search Strategy #	Search Term
1	“Process utility”
2	"treatment related attributes" AND (utilities OR "utility measurement")
3	"dose frequency" AND (utilities OR "utility measurement")
4	"drug administration" AND (utilities OR "utility measurement")

All searches were performed in PubMed, no date restrictions were applied. Searches were initially performed in April 2013, and updated in August 2015. The following inclusion/exclusion criteria were applied:

Inclusion Criteria

- Studies published in English language;
- Reporting an empirical estimate of process utility which can be incorporated into the QALY calculation.
- Chained conjoint analysis studies.

Exclusion Criteria

- Studies in language other than English;
- Economic models which were not the primary utility elicitation study for process utility estimates;
- Studies which do not report process utility estimates;
- Standard conjoint analysis studies and willingness to pay studies. These were excluded as they were unable to provide a utility estimate on a scale anchored by 0 representing dead, and 1 representing full health. Where a chained conjoint analysis approach was taken, and empirical measures of process utility were reported, these studies were included.

The study selection process was performed by one researcher who initially screened the titles and abstracts for eligibility. The full text versions of all included studies were then obtained and reviewed for eligibility by one researcher using the same inclusion and exclusion criteria. Any uncertainty about the inclusion of studies was checked by a second researcher. The final list of included studies was sent to experts in the field (Andrew Lloyd, Phil Shackley), to ask whether they were aware of additional studies, not identified by the searches. Data were extracted by one researcher into data extraction tables in Microsoft word (Tables 10-12). Any uncertainty surrounding the extraction of data was checked with a second researcher.

4.2.3 Results

The literature search results are presented in Figure 14. The PubMed searches yielded a total of 118 titles and abstracts. Twenty-six of these were identified through search strategy 1 and had therefore embraced the term “process utility”. The remainder used the alternative statements such as “treatment related attribute”. An additional 9 studies were identified through a hand search of the reference lists. After removing 3 duplicates, 124 articles were screened for eligibility. Studies were excluded for the following reasons: clinical paper (n = 20); study not related to treatment attributes (n = 27); review (n = 4); no outcomes of interest reported (n = 47); methodological paper (n = 9). An additional 2 were identified through the supervisors (SD) knowledge of the field. A total of 20 studies met the inclusion criteria and were included in the review. The update to the searches in August 2015 also identified 4 literature reviews into

aspects of process utility, one of which was the publication from this review (Brennan & Dixon, 2013). The 4 reviews will also be compared and contrasted within this chapter.

Cost-utility Compatible Measures of Process Utility.

In total, 20 studies were identified which provided empirical measures of process utility which could be incorporated into the QALY calculation. A summary of the study characteristics can be seen in Table 10.

Figure 14. PRISMA Diagram

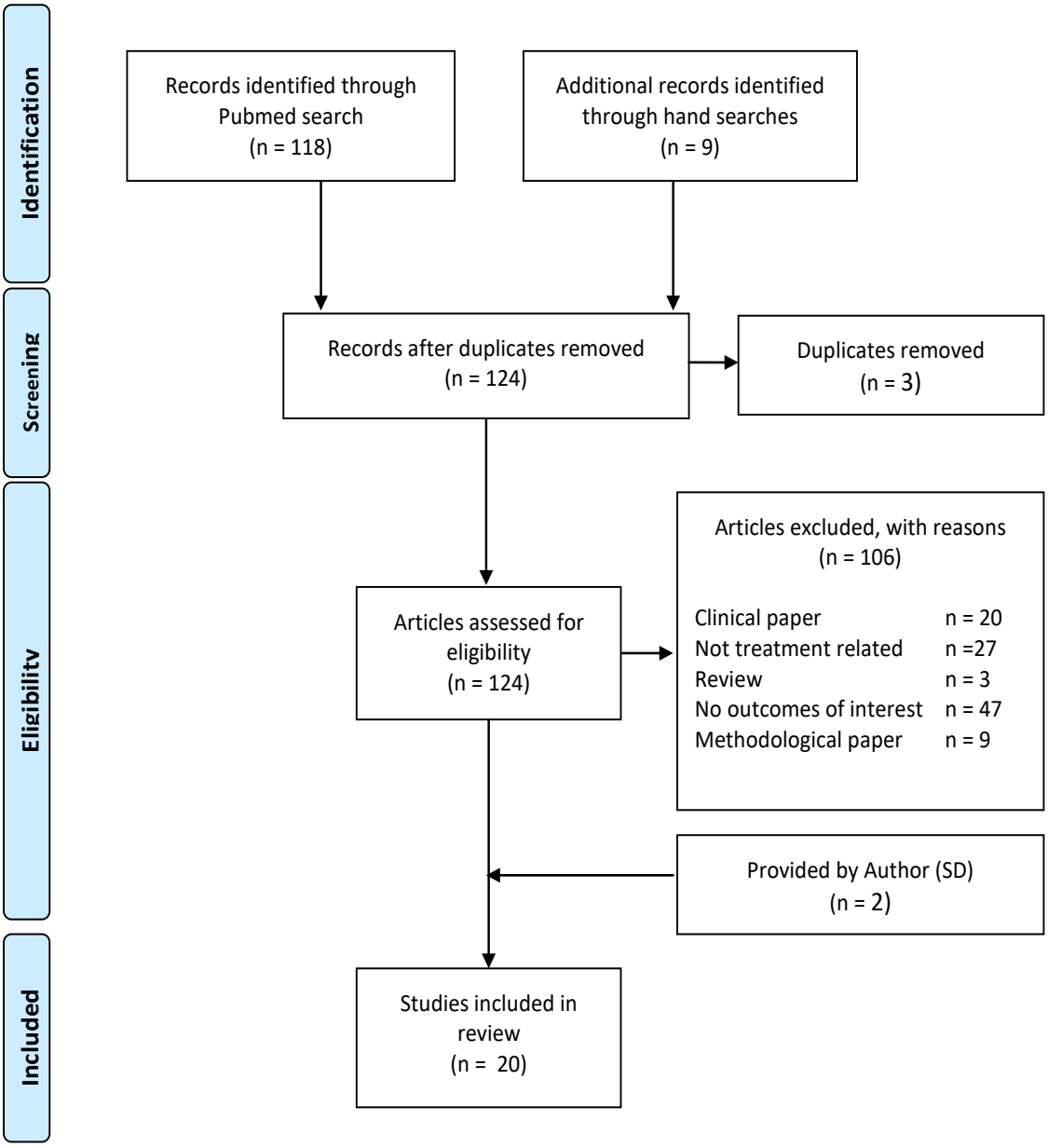


Table 10. Summary of cost-utility compatible studies.

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
Birch (2003)	USA	Cervical screening	Medical management of mildly abnormal Pap smears	Watchful wait (conservative) vs. aggressive management	6 scenarios: 3 for conservative management and 3 for aggressive	Patients (women)	Two stage SG Upper anchor: full health Lower anchor: invasive cancer requiring hysterectomy
Boye (2011)	Scotland	T2DM treatment	Injectable treatments	Flexible dosing, number of injections	9 scenarios: 1 basic health state, and 8 using basic health state plus statements describing	Patients	Two stage SG Upper anchor: perfect health Lower anchor: worst health represented by a person with T2DM,

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
					injectable treatment added to oral treatment		blood glucose levels not in control, body weight 10% higher than current weight and suffering from multiple complications
Cairns (1996)	Scotland	Antenatal screening for cystic fibrosis	Stepwise vs. couple carrier screening	Level of information	6 scenarios: 1 with 100% chance the foetus had cystic fibrosis; 5 including different levels of information being available to the couple	General population	SG Upper anchor: full information Lower anchor: foetal loss
Chancellor	UK	T2DM treatment	Inhaled vs.	Route of	5 scenarios: two for	Patients	Two stage TTO and

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
(2008)			injectable insulin regimens	administration (inhaled vs. injectable)	T1DM (injectable and inhaled); 2 for T2DM (injectable and inhaled)		EQ-5D Upper anchor: full health Lower anchor: least preferred health state
Cook (2007)[	Australia	Gallstone disease treatment	Open vs. laparoscopic surgery and ESWL	Open vs. laparoscopic surgery	7 model health states, 2 related to laparoscopic vs. open cholecystectomy	Patients	Two stage TTO Upper anchor: normal health Lower anchor: worst temporary health state
Evans (2013)	UK, Canada, Sweden	Diabetes	Attributes associated with insulin	Dose timing, frequent injections and	6 health states: : once-daily flexible injection, once daily	General population, and sub	TTO SG

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
			treatment.	self-measured blood glucose (SMBG)	fixed injection and 2 fixed injections. At 2 treatment intensities: Basal only Basal and bolus	analyses on patients with ● T1DM ● T2DM	Both approaches were modified to consist of 1 multiple choice question.
Howard (2008)	Australia	Cervical screening	Medical management of mildly abnormal Pap smears	Repeat Pap smear vs. immediate HPV test	4 clinical management scenarios: 1 anchor state for cervical cancer needing hysterectomy; 2 describing management of a lesion that resolved	General population (women)	Two stage SG Upper anchor: full health Lower anchor: cancer state needing treatment with hysterectomy

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
					spontaneously; 2 describing scenarios that required treatment		
Hutchins (2015)	US	Cardiovascular prevention	Daily pills for cardiovascular prevention	Dose frequency	Asked questions about respondents' medication taking regimens, and used multiple choice questions to	General population; internet sample.	TTO[1], SG[2]

[1] Utility derived by dividing the maximum amount of time each respondent was willing to give up at the end of their life to avoid having to take a medication (determined via a single multiple-choice question), by the mean amount of time each respondent had remaining in their life (with average life expectancy 78 years) and subtracting it from 1.

[2] Utility calculated as 1 minus the maximum risk of death the participant was willing to accept to avoid taking medication daily, assessed via a single multiple-choice question. Potential responses included 0,1:10:100:1,000:10,000 or 100,000 with a constant denominator of 1,000,000.

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
					determine TTO and SG estimates.		
Johnson (1996)	USA	AIDS-related cytomegalovirus retinitis treatment	Intravenous ganciclovir compared with oral ganciclovir	Route of administration (intravenous vs. oral)	3 health states: patients current state of health; IV therapy: oral therapy	Patients	TTO
Kauf (2008)	Multi-country (Europe and the USA)	HIV treatment	Treatment regimen attributes	Dose frequency, symmetrical regimen, treatment taken with food/drink/empty stomach	No, SF-36 in HIV patients	Patients	SF-6D

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
Matza (2013)	UK	Bone metastasis	Intravenous (infusion) (with renal monitoring) versus subcutaneous injection (without renal monitoring)		Hypothetical health states: bone metastases; plus 9 additional stated describing treatment involving either infusion or injection	General population	VAS and TTO
Osborne (2007)	Australia	Haematology treatment	Oral vs. subcutaneous iron chelation	Route of administration	3 health states: anchor state with patient with iron chelation no description of treatment; anchor	General population	TTO

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
					state plus treatment via subcutaneous injection; anchor state plus treatment via oral medication		
Osborne (2012)	Australia	Schizophrenia	Long-acting injectable antipsychotic treatment	Dose-frequency	4 vignettes. One relapsed/untreated schizophrenia; 3 standardised picture of well-managed schizophrenia with variations in	General population	TTO

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
					intervals between injections – 2 weeks, 4 weeks, 3 months.		
Polster (2010)	USA	T2DM treatment	GLP-1 products- liraglutide and exenatide	Dosing frequency (QD vs. BID)	4 product profiles	Patients	TTO and CA Upper anchor: life free of diabetes
Sadigh (2013)	USA	Diagnosis of coronary artery disease	Coronary computed tomography angiography and conventional catheter	Administration type	Telephone interviews with patients, plus descriptions of hypothetical tests.	Patients	WTO and rating scale

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
			angiography				
Salkeld (2004)	Australia	Preventative health care	Hip protectors vs. no hip protectors	Comfort, reassurance	6 scenarios: full health; 2 process states; 3 psychological states	General population (women)	TTO
Schmier (2002)	Canada, Australia, and USA	Chronic pain treatment	Opioid treatment, oral vs. patch	Route of administration	22 Health states characterised by treatment attributes	Patients	Conjoint analysis
Swan (2003)	USA	Cerebrovascular disease screening	Diagnostic testing	Conventional angiography vs. MR angiography	Chronic and short-term health states	Patients	WTO
Swan (2006)	USA	Breast cancer screening	Breast biopsy	CNB vs. EXB	Chronic and short-term health states	Patients	WTO

Author (year)	Country	Indication	Interventions	Source of process utility	Hypothetical health states	Population	Valuation methodology
Swan (2000)	USA	Peripheral vascular disease screening	Diagnostic testing	MRI vs. conventional angiography	Chronic and short-term health states	Patients	WTO

BID: twice a day; CA: conjoint analysis; CNB: core needle biopsy; EQ-5D: EuroQol 5 Dimensions; ESWL: extracorporeal shockwave lithotripsy; EXB: excisional surgical biopsy; GLP-1: glucagon-like peptide; HPV: human papilloma virus; IV: intravenous; MR: magnetic resonance; QD: once daily; SF-36: Short Form 36; SG: standard gamble ; T1DM: type 1 diabetes mellitus;T2DM: type 2 diabetes mellitus; TTO: time tradeoff; USA: United States of America; WTO wait tradeoff.

The earliest of the included studies was performed by Cook et al. (1994) who undertook a CUA comparing extracorporeal shock wave lithotripsy (ESWL) to open and laparoscopic cholecystectomy for gallstone disease. ESWL is a non-invasive day-case procedure with minor levels of post-procedure morbidity, but diarrhoea and nausea are experienced by some patients. A laparoscopic cholecystectomy is a minor operation, and patients can usually return to normal activities 2 to 3 days post-operatively. There is a small probability of serious complications; however, post-operative morbidity is significantly reduced when compared to open surgery. Respondents were asked to value hypothetical health states using a two-stage SG technique. The author's postulate that the short duration of the disutility experienced post surgery may be overshadowed by the disutility arising from the process of receiving surgery and therefore having the associated risk of death. To overcome this they asked respondents to quantify disutility for the operation (ex ante) and the riskless health state (ex post). The results were summed to provide a "partial ex ante" perspective. In essence the ex ante evaluation of the operation represents the "process" of receiving care, and the ex post evaluation of the health states refers to the health outcomes.

The authors calculated a QALY loss associated with the three procedures from both an ex post perspective, and an ex ante perspective. The results indicated that the ex ante approach lead to considerably higher QALY loss. The cost per QALY indicated that the open cholecystectomy was always dominated, but that the

choice of approach (ex ante vs ex post) had a significant impact on the cost-effectiveness of the laparoscopic cholecystectomy and ESWL. Where an ex post approach was taken laparoscopic cholecystectomy dominated. When an ex ante approach was used the QALY loss from laparoscopic cholecystectomy was greater than the loss associated with ESWL. This study clearly indicates that the choice of components included in the patients' utility function can have a significant impact on the overall cost-effectiveness results.

Schmier et al. (2002) explored process utility associated with the treatment of chronic pain in patients receiving opioid analgesia. They used conjoint analysis techniques to explore patient preferences for four key treatment attributes: pain control; side effects; severity of side effects; and the route of administration (oral versus patch). The latter of which refers to a "process" of health care. The study included 96 patients with non-malignant pain and 25 patients with malignant pain. Patients with malignancy were also asked to complete the QLQ-C30 and patients with non-malignant pain were asked to complete the MOS SF-36. A computer programme was used to convert participants' responses to utility estimates. The results showed slight differences in the mean utilities assigned to health states that differed only by oral versus patch route of administration. For patients with non-malignant pain there was a small and consistent preference for the oral

administration of opioids. For patients with malignant pain preferences differed by severity of side effects, with patients preferring oral administration when pain was mild and patch administration when pain was severe. Utility differences between the same health states but receiving medication by oral versus patch medication ranged from no difference (for severe respiratory depression health state) to 0.16 in preference for oral medication for the health state describing moderate respiratory depression.

Osborne et al. (2007) conducted a time trade-off (TTO) study to explore the utility associated with subcutaneous infusion when compared to the oral administration of iron chelation therapy. They used TTO techniques to estimate preferences for three hypothetical health states with differing treatment modalities but identical treatment outcomes: an anchor state where the patient has iron chelation but the treatment itself is not described; the anchor state plus treatment via a subcutaneous infusion; the anchor state plus treatment via once-daily oral treatment. They found a mean difference of 0.23 (0.27) in the utility estimates for the two treatment health states in preference for oral administration. This was found to be statistically significant.

A further study by Osborne et al. (2012) used time trade-off methodology to explore utility differences associated with long-acting injectable treatment for schizophrenia. Vignettes were developed: the first represented relapsed/untreated schizophrenia; the remaining 3 represented well-managed schizophrenia with variations in the intervals between injections (once every 2 weeks, 4 weeks and 3 months). Ninety-eight members of the general population completed the TTO interview. The results showed that utility values increased as the interval between administrations increased. With the 2 week interval resulting in a mean utility of 0.61, 4 weekly intervals, resulting in a mean utility of 0.65 and 3 month intervals resulting in a 0.70 utility estimate.

Kauf et al. (2008) used SF-36 responses from five clinical trials in HIV to explore the marginal impact on utilities of different dosing frequencies in HAART treatment in patients with HIV. Dosing frequencies were categorised as one or two times per day, with twice per day (BID) as the reference level. Food and drink requirements were also explored (whether medication must be taken with food and/or drink or on an empty stomach). A total of 1327 participants were included in the study. Small differences between utility estimates according to dosing strategies were identified: for medication which had to be taken with food -0.028; medications to be taken on an empty stomach -0.001; medications taken once per day 0.020.

Hutchins et al. (2015) also estimated utilities for dosing frequency. They completed an internet based survey of 1,000 US residents which used modified TTO (primary analysis), SG and WTP (secondary analyses) to derive utility estimates. The mean utility decrement for taking 1 pill per day (compared to no pills per day) was 0.01. Respondents were told the question related to taking a pill to prevent a heart attack or stroke, and they should assume that by taking the pill they will live their life free from both. TTO estimates were also derived for taking 2 pills daily, and 1 pill twice daily. However, the decrement remained at 0.01.

A further study by Evans et al. (2013) used TTO to estimate utilities associated with different dosing regimens in patients with diabetes. They used hypothetical health states to compare a once daily flexible injection; a once daily fixed injection and 2 fixed injections. These 3 dosing strategies were compared using 2 treatment intensities: basal only, where respondents were asked to imagine having baseline diabetes and administer basal insulin injections (background insulin, long acting); basal and bolus, where they were also taking a bolus insulin injection with every meal (short acting insulin, needed at meal times). The study was based on a general population sample with sub-groups of patients with T1DM and T2DM. Within their TTO estimations they used time horizons which were based on each respondents projected life expectancy. Utility estimates ranged from 0.004 for comparisons of

one flexible injection versus one fixed injection in the sample with T1DM and T2DM patients, to 0.055 in comparisons of one flexible injection to 2 fixed injections.

Johnson et al. (1996) compared utilities for intravenous (IV) medication compared with oral medication for the treatment of AIDS-related cytomegalovirus retinitis. They used TTO techniques to value hypothetical health states in 80 patients. They found that 75% of patients preferred oral medication, and identified a difference in median utilities of 0.362 and mean utilities of 0.223.

Matza et al. (2013) also used TTO and compared IV medication with subcutaneous injection in patients with bone metastasis. They developed 10 hypothetical health states: one represented a basic health state for patients with cancer and bone metastasis but no treatment description. A further 9 health states described the treatments involving either injection or infusion. They purposefully kept all treatment efficacy constant, so the only differences were associated with the treatment modalities. They were asked to consider remaining in the health states for 2 years. Respondents were first asked to rate the health states using the VAS, followed by TTO questions. They found differences in utility between the treatment modalities, with injections being preferred over infusions.

Several studies have explored process utility associated with treatment attributes in patients with diabetes. Chancellor et al. (2008) explored patient preferences for inhaled versus injectable insulin regimens. Written descriptions were developed for five clinical scenarios (two for type 1 and two for type 2 diabetes), with two alternative insulin scenarios described for each: injectable only or inhaled insulin to replace or reduce the number of daily injections. Equal efficacy was assumed within each scenario pair. 344 patients participated in the study, with each rating scenario pairs associated with their diabetes type, and their current health using both TTO techniques and the EQ-5D. A majority of patients preferred inhalation to injectable treatment. Differences in utilities in favour of inhaled administration ranged from 0.04 to 0.09 using TTO techniques, and 0.02 to 0.07 using the EQ-5D.

Polster et al. (2010) compared patients' preferences for attributes of two diabetes treatments. These attributes included comparisons of once daily versus twice daily doses. They used a survey instrument consisting of the EQ-5D, a TTO exercise where respondents compared the profiles of the two treatments, and a conjoint exercise in which respondents compared a series of TTO exercises for hypothetical product profiles. This combination of TTO and conjoint analysis methods allowed utility estimates to be derived on the standardised utility scale anchored by 0 (dead) and 1 (perfect health). Although, within the TTO task the authors chose to

use “life free of diabetes” in place of the anchor “perfect health” to more closely reflect a realistic decision scenario for respondents, and to limit the impact of potential co-morbidities. The results showed a 0.005 increase in utility for movement from a twice daily dose frequency to a once a day frequency. Studies which used conjoint analysis alone would not result in estimates of utility, and therefore were excluded from the review.

Boye et al. (2011) explored the differences in utility associated with three injection-related attributes in patients with diabetes. Two of these were related to process: dose frequency and dose flexibility. They asked patients with type 2 diabetes to value hypothetical health states and their own current health using a two stage SG technique. A total of 151 patients completed interviews. They found that weekly injections were associated with an average added utility of 0.023 when compared to daily injections. Dose flexibility compared medication which can be taken at any time of the day to medication taken with food. They found an added average utility of 0.006 for medication taken at any time of the day.

Studies have also been undertaken to evaluate process utility associated with screening/testing procedures. Birch et al. (2003) compared aggressive versus conservative management approaches to the follow-up of mildly abnormal Pap

smears. They asked women attending family planning clinics to evaluate 6 hypothetical scenarios which differed in management approach but had the same health outcomes using a two-stage SG technique. They found that for comparisons where no abnormalities were detected, an immediate colposcopy (normal) versus observation followed by 3 repeat smears compared to an immediate colposcopy resulted in an increase in utility of 0.031. For comparisons where mild abnormalities were found, they compared 3 repeat smears followed by cryotherapy to an immediate colposcopy followed by cryotherapy, and identified an associated -0.021 difference in mean utilities. For comparisons where abnormalities were found which necessitated a cone biopsy, they compared 3 repeat smears followed by biopsy to an immediate colposcopy followed by biopsy and identified a -0.017 difference in utility estimates.

Howard et al. (2012) also explored process utility associated with alternative approaches to the testing and management of abnormal atypical squamous cells in women. They asked women to evaluate four clinical management scenarios, using two-stage SG. The scenarios either described the management of a lesion that resolved spontaneously following either a repeat Pap smear or negative HPV test; or scenarios describing the management of a lesion that required treatment following either a repeat Pap smear or a positive Pap test. The results found that

there was evidence of process utility, with HPV testing strategies having lower valuations than repeat Pap smear tests where the clinical outcome was the same. These results indicate that including measurements of process utility in studies would have implications on management decisions, as although it is suggested that HPV triage testing can offer a more effective management strategy over conventional management for women with atypical squamous cells, the women themselves do not necessarily see HPV testing as having higher utility than conventional management with repeat smear.

Cairns et al. (1996) used SG techniques to estimate the value placed on improved information in antenatal screening for cystic fibrosis carrier status. Two methods of administering the cystic fibrosis carrier status screening test were evaluated; stepwise screening and couple screening. Stepwise screening was initially offered to the pregnant woman, only if the test is positive was the test offered to her partner. In couple screening both partners were tested at the same time, and the result of the test combined the outcome. If they had a positive test, they are then offered prenatal diagnosis, where the cystic fibrosis status of the foetus was determined. This incurred approximately a 1 in 200 risk of foetus loss. Women were asked to trade off the risk of foetal loss with improved information from diagnostic testing. The women were initially presented with the scenario where

there was a 1 in 100 chance that the foetus had cystic fibrosis. For the remaining 5 scenarios the SG methods were used to establish the risk of foetal loss associated with amniocentesis at which the woman was indifferent as to whether or not the diagnostic test was performed. The results of the study indicated that as the risk of detecting a cystic fibrosis infected foetus increased, the associated utility estimates decreased. For both stepwise screening and couple screening, the expected utility of screening was found to be greater than that of not screening, and the expected utility of couple screening was found to be similar. The study found no statistically significant difference between the expected utilities of the two methods, however, at a patient level, stepwise screening was preferred; couple screening was preferred when considering only those women who preferred screening to no screening.

Swan et al. (2000) used a variation of the TTO methodology to assess patient's preferences for magnetic resonance angiography (MRA) versus x-ray angiography (XRA) in patients with peripheral vascular disease (PVD). They use wait trade off techniques (WTO) by asking the patient to trade off extended time with the condition being diagnosed in order to avoid the adverse effects of one screening text compared to another. They found significant differences between patients' wait times and corresponding disutilities for MRA versus XRA which resulted in a

difference of 9.36 quality adjusted life days (QALD) between tests and equates to a 0.026 difference in utility estimates.

Swan and Sainfort (2003) used the same WTO methodology to explore process utility associated with imaging in cerebrovascular disease. They compared the preferences of 89 patients for conventional versus magnetic resonance (MR). Patients had experienced both tests. They calculated a weighted difference between the time a patient was willing to wait for a test result and treatment after a hypothetical "ideal" test and the choice to undergo conventional angiography or MR angiography with immediate treatment. The study found a significant difference between patient preferences for conventional and MR angiography in favour of MR, with a utility increment of 0.014.

Swan et al. (2006) again used WTO techniques to investigate process utility associated with different methods of biopsy for the detection of breast cancer. They compared preferences for core-needle biopsy (CNB) with the more invasive excisional surgical biopsy (EXB) in 38 women. The study indicated that women were willing to wait for significantly longer to avoid the EXB. They used patient's baseline and test-related anxiety, measured using TTO, to scale the WTO. Median TTO

preferences for baseline was 1.00 and anxiety was 0.93. The QALD toll for EXB was 1.5 and for CNB was 0.5, this equates to a utility value of 0.001 in favour of CNB.

Sadigh et al. (2015) also used WTO methodology to compare patients preferences for computed tomography angioplasty (CCTA) versus conventional catheter angiography (CCA) for the diagnosis of coronary artery disease (CAD). Baseline utility values for a patient with CAD were derived using SG. WTO was used to elicit utilities for each intervention. The disutility associated with the CCTA was 2.16 QALDs and for the CAA was 1.30 QALDs. This resulted in a QALD difference of 0.86, and a utility value of 0.0024 in favour of CAA.

The final included study identified in this review, explored the presence of process utility in the prevention of hip fractures. Salkeld et al. (2004) asked participants to value 6 scenarios using TTO. The scenarios included a “full health” state, two process states (wearing hip protectors and feeling discomfort; wearing hip protectors and feeling confident), and three states associated with psychological outcomes of regret and elation. Utility estimates for the process states were 0.31 for wearing hip protectors and feeling discomfort compared to 0.64 for wearing hip protectors and feeling confident. Estimates for the health state with elation for a

fall but no hip fracture was 0.63; for a fall, hip fracture and regret was 0.10 and for a fall with no hip fracture and elation was 0.71.

Comparison of Key Methodological Characteristics

Hypothetical Health States

Seventeen of the 20 included studies used hypothetical health states to value process utility. Whereby, health states were developed which described not only health outcomes, but also considered process outcomes. The remaining 3 used alternative approaches: Sadigh et al. (2013), in their study into patient preferences for coronary computed tomography (CT) angiography versus conventional catheter angiography, identified 30 patients who had previously received both tests. They then conducted telephone interviews in which they asked respondents about their symptoms, quality of life, and experience in preparing for the different diagnostic tests. They were then asked about their preference for undergoing a CT or catheter angiography and immediate treatment versus having waiting for a hypothetical test with no side effects (an "ideal" test) and treatment. This information was used to calculate the disutility for each test.

Kauf et al. (2008) applied regression techniques to SF-6D data collected from patients within clinical trials, where additional clinical data was also available,

including frequency of drug administration. They used this to determine whether differences in administration frequency influenced utility estimates. This therefore negated the need for hypothetical health states. The final study by Hutchins et al. (2015) questioned respondents via telephone interviews on their medication taking regimens and used multiple choice questions to determine TTO and SG estimates of process utility.

Methodological Approaches

A variety of approaches were used to detect and measure PU, these are summarised in Figure 15. The most commonly used approach was TTO, which was used by 8 of the included studies. As is commonly the case in TTO studies, here were some variations on the standard TTO approach (Attema et al. 2013). Polster et al. (2010) combined the use of TTO with conjoint analysis to value the profiles of 2 diabetes treatments, which varied in terms of dosing, control of blood sugar, incidence of nausea and incidence of hypoglycaemia. They initially performed a TTO exercise where respondents were presented with a pair of product profiles. Conjoint analysis was then used to establish the relative importance of 4 attributes: efficacy; dosing schedule; nausea; hypoglycaemia. Hutchins et al. (2015) also used a modified TTO approach. Rather than using the traditional approach where open-

ended questions are used in an interview format to “ping pong” the time willing to trade until a point of indifference is met, the TTO values were derived using close ended questions in a multiple-choice format. This allowed the survey to be administered online and to a large sample. Their calculation involved dividing the maximum amount of time each respondent was willing to give up at the end of their life to avoid having to take the medication (determined through a single multiple choice question) by the mean amount of time each respondent had remaining in his or her life (based on average life expectancy of 78 years), and subtracting it from 1. The potential responses for the question included none, 1 week, 2 weeks, 1 month, 3 months, 6 months, 9 months, 12 months, 18 months or 24 months.

Hutchins et al. (2015) also used a modified SG question, again to allow it to be completed online and to a large sample. The SG estimate was calculated as 1 minus the maximum risk of death the participant was willing to accept to avoid taking medication daily. It was again assessed through a single multiple-choice question. Potential responses included 0, 1, 10, 100, 1000, 10 000, or 100 000, with a constant denominator of 1 000 000. This resulted in a mean utility decrement of 0.009 for moving from 0 tablets per day to 1. The same description of the medication was provided. Although, as the authors point out, their estimates of

disutility associated with the frequency of pill taking are in line with previous studies. The impact of their decision to modify the SG and TTO techniques on the utility estimates derived is unknown. Due to the large sample size Hutchins et al. (2015) also provided estimates by participant characteristics. SG techniques were used in a total of 6 studies.

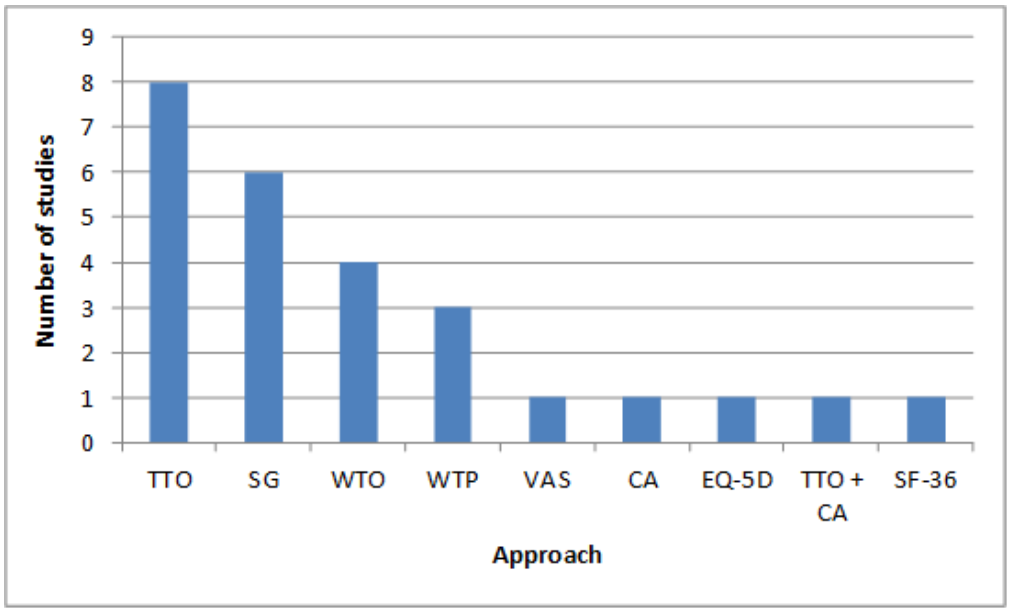
A further study which used conjoint analysis was Schmier et al. (2002) who calculated utility for patients receiving pain treatment in patients with and without cancer. Following completion of a CA exercise, regression analysis was used to estimate utilities.

WTO was used in 3 studies, and can be used to provide estimates which can be included in the QALY calculation; this advantage was the rationale for its selection by Swan who lead all 3 publications (Swan et al. 2000; 2003; 2006). This approach is a modification on TTO in which the respondent trade's time waiting with symptoms for an "ideal" test, compared to undergoing a traumatic test followed by immediate treatment.

Two studies used multi-attribute preference measures: Kauf et al. (2008) used SF-36 data collected from previously completed Phase II, III and IV clinical studies, alongside additional clinical data such as the number of tablets taken per day and

whether treatment must be taken with food or on an empty stomach. Regression analysis was used to determine utility differences according to these process parameters. The study found utility reduced with the requirement to take with food or on an empty stomach, and increased if administered orally as opposed to intravenously or if it was taken once per day instead of twice per day. However, none of these differences in utility reached statistical significance. Chancellor et al. (2006) used the EQ-5D alongside TTO to obtain 2 sets of utility estimates, and found estimates derived using the EQ-5D to be lower than those derived from TTO.

Figure 15. Summary of Methodological Approaches.



Anchoring of Death

Several of the included studies used a two-stage or chaining technique for the valuation of health states, specifically when the states under evaluation did not contain a risk of death (Matza et al. 2013; Boye et al. 2011; Chancellor et al. 2008; Birch et al. 2003; Howard et al. 2012; Cairns et al. 2011). Using this methodology, rather than immediate death being compared to full health, an "anchor" state (which is better than death but worse than the scenarios under evaluation) is used. This anchor state is then compared with an intermediate health state of interest. Scores are then converted onto a scale of 0 representing death and 1 representing full health. This is the methodology suggested by Torrance (1986) for the valuation of temporary health states. For some studies the exact methodology used is not clear, for example Hutchins et al. (2015) report that utilities "usually range from 0 to 1, with 0 representing death and 1 representing perfect health", however, in their "modified" approaches to TTO and SG, it is not clear whether their estimates were anchored on this scale. There was also a lack of clarity in the exact TTO procedure used by Evans et al. (2013) in their study into flexible dosing for insulin therapy, where no description of the rescaling of results on a scale anchored by 0 (dead) and 1 (full health) is reported. This lack of clarity in reporting makes it difficult for us to fully comprehend the utility estimates derived.

Utility Estimates

All studies identified the presence of process utility. Table 11 shows the results of studies evaluating process utility associated with treatments. Measures of process utility for different drug delivery methods ranged from 0.004 (Matza et al. 2013) to 0.27 (Osborne et al. 2007). Utility estimates associated with different dosing strategies ranged from 0.004 (Evans et al. 2013) to 0.09 (Osbourne et al. 2012). Estimates for convenience (able to take on an empty stomach) ranged from 0.001 to 0.028 (Kauf et al. 2008).

Estimates of process utility associated with screening and testing procedures can be seen in Table 12. Measures ranged from 0.0005 (Howard et al. 2008) to 0.031 (Birch et al. 2003). Both of these estimates were obtained for different management approaches to cervical cancer screening.

Table 11. Process utility gains from drug delivery methods/dosing strategies.

Author (year)	Source of process utility	Process utility gained
Boye (2011)	Daily injections to weekly injections	0.023
	To be taken with food or empty stomach to flexible dosage	0.006
Chancellor (2008)	Injectable to inhaled medication (TTO)	0.04 to 0.09
	Injectable to inhaled medication (EQ-5D)	0.02 to 0.07
Cook (1994)	From ESWL to open cholecystectomy	–0.045
	From ESWL to laparoscopic cholecystectomy	–0.001
	From open to laparoscopic cholecystectomy	0.044
Evans (2013)	General population sample: From one flexible injection to one fixed injection (basal[1])	0.016
	From one flexible injection to one fixed injection (basal-bolus[1])	0.013

[1] Basal dosing has to be taken at the same time every day/twice a day.

Author (year)	Source of process utility	Process utility gained
	From one fixed injection to 2 fixed injection (basal)	0.039
	From one fixed injection to 2 fixed injection (basal-bolus)	0.022
	From one 1 flexible injection to 2 fixed injection (basal)	0.055
	From one 1 flexible injection to 2 fixed injections (basal-bolus)	0.036
	T2DM sample	
	1 flexible injection to 1 fixed timed injection (basal)	0.015
	From one fixed injection to 2 fixed injection (basal)	0.042
	From one 1 flexible injection to 2 fixed injection (basal)	0.057

[1] Some patients requiring treatment intensification, have frequent injections of fast-acting bolus (meal-time) insulin in a basal-bolus regimen.

Author (year)	Source of process utility	Process utility gained
	T1DM & T2DM sample	
	1 flexible injection to 1 fixed timed injection (basal-bolus)	0.004
	From one fixed injection to 2 fixed injection (basal-bolus)	0.021
	From one 1 flexible injection to 2 fixed injection (basal-bolus)	0.025
Hutchins (2015)	No pills per day to 2 pills per day (TTO)	-0.01
	No pills per day to 1 pill twice daily (SG)	- 0.000909
Johnson (1996)	From IV to oral drug administration	0.223
Kauf (2008)	Medication to be taken with food	–0.028
	Medication to be taken on empty stomach	–0.001
	Medication taken twice per day to once per day	0.020

Author (year)	Source of process utility	Process utility gained
Matza (2013)[1]	Injection	-0.004
	30 minute infusion	-0.023
	2 hour infusion	-0.037
	30 minute infusion + renal monitoring (blood drawn on same day as infusion)	-0.050
	30 minute infusion + renal monitoring (blood 2 days prior to infusion)	-0.066
	Chemotherapy	-0.175
	Injection + chemotherapy	-0.190
	30-minute infusion + chemotherapy	-0.202
	2 hour infusion + chemotherapy	-0.211

[1] The disutility for each treatment modality represents the impact of adding each treatment modality to an otherwise identical health state (Matza et al., 2013. p861).

Author (year)	Source of process utility	Process utility gained
Osborne (2007)	Subcutaneous administration to oral administration	0.27
Osbourne (2012)	Injections every 2 weeks to injections every 3 weeks	0.04
	Injections every 2 weeks to injections every 3 months	0.09
	Injections every 3 weeks to injections every 3 months	0.05
Polster (2010)	Twice a day dose frequency to once a day	0.005

EQ-5D: EuroQol 5 Dimensions; ESWL: extracorporeal shock wave lithotripsy; IV: intravenous; TTO: time tradeoff.

Table 12. Process utility gains from changing screening/testing procedures

Author (year)	Source of process utility	Process utility gained
Birch (2003)	Immediate colposcopy (all normal) to Observation/3 repeat smears (all normal) for mildly abnormal Pap smear	0.031
	3 repeat Pap smear all mildly abnormal to immediate colposcopy all mildly abnormal for mildly abnormal Pap smear	–0.022
	3 repeat Pap smear, immediate colposcopy and biopsy (abnormal), cone biopsy (moderately abnormal), repeat 3 Pap smears to colposcopy and biopsy (abnormal), colposcopy and biopsy (normal) repeat Pap smears	0.016
Cairns (1996)	Stepwise CF screening	0.00075465
	Couple CF screening	0.00072325
Howard (2008)	repeat Pap smear, spontaneous resolution/HPV negative to immediate HPV test, spontaneous resolution/HPV negative	0.0005
	repeat Pap smear, abnormal pap/HPV positive then treatment to immediate HPV test, abnormal pap/HPV positive then treatment	0.0302

Author (year)	Source of process utility	Process utility gained
Sadigh (2015)	Computed Tomography Angiography to conventional catheter angioplasty	0.0024
Swan (2006)	Excisional needle biopsy to core needle biopsy	0.001
Swan (2003	Conventional magnetic resonance to magnetic resonance angiography	0.014
Swan (2000)	XRA to MRX	0.026

CF: cystic fibrosis; HPV: human papilloma virus; MRA: magnetic resonance angiography; XRA: x-ray angiography.

Literature Reviews

The searches identified 4 literature reviews into aspects of process utility. One of which stemmed from the review above, prior to the update in 2015, and focused specifically on the studies where utility estimates can be incorporated into the QALY (Brennan & Dixon 2013). The paper included 15 studies published between 1996 and 2012. Of these 8 explored process utility associated with treatments; 6 explored process utility associated with screening procedures of tests; one was performed in preventative care. A variety of approaches were used including SG, TTO, CA, WTO. The estimates of utility varied considerably between studies. For example, estimates for convenience ranged from 0.001 to 0.028, estimates of process utility for screening and testing procedures ranged from 0.0005 to 0.031, and estimates for different drug delivery methods ranged from 0.02 to 0.27. We concluded that the identification of studies was difficult given the lack of consistency in reporting. However, from those studies identified the evidence does support the existence of process utility, although given the disparity between estimates and methodological approaches used it is difficult to know to what extent.

Opmeer et al, (2010) suggest that health outcomes may not be the only outcomes which are important to consider within health care, but that attributes related to the process of receiving care are also important. They therefore performed a

literature review to assess the extent to which valuations of process and non-health outcomes had been undertaken, in which specialities and using which techniques. Their study did not focus on QALY compatible utilities. Their wider inclusion criteria led to a larger review with 567 studies being included for analysis. They identified studies across many medical fields, with most being performed in treatments associated with oncology, and others including gynaecology, pulmonary, cardiology, gastroenterology and infectious disease.

Preferences for processes were elicited using different methods, most commonly using a binary choice (A vs B) (45%) or judgement of individual attributes (15%) or trade-off questions (13%). It is not clear if the latter related specifically to the TTO procedure introduced by Torrance (1986), it may also include variations of this such as WTO. Their study focused on the number of studies which consider process rather than any measures of process and specific estimates.

They also identified that the number of studies which consider process has steadily increased between 1980 and 2002, and rose sharply from 2003 to 2006. Although the focus of the review differed to our publication, some of the findings were similar: they suggested that the consideration of none health attributes is becoming increasingly important and prevalent in health care research, and

concluded that the identification of studies was made harder by the lack of standardised terminology, but that this lack of consistent terminology also hindered data extraction, for example in the classification of non health outcomes.

Ryan et al. (2014) undertook a study concentrating on the evaluation of patient's experiences of receiving health care. They build on a literature review by Entwistle et al. (2012) which asked what experiences of health care matter to patients and why, and produced a "concept map" which reports important capabilities which people expect health care services and staff to have, and the ways they enable patients. They identified 3 groups of responses: having characteristics that equip them to deliver consistently good care; acting in ways that show they are willing and competent; enabling patients to be and do what they value being and doing. They therefore undertook a further literature review to understand the application of economic methods to studies incorporating a measure of patient's experiences of receiving care. They began by identifying the economic methods which could be applied, and then performed a search to identify applications of these methods to valuing experiences of receiving care. A systematic review was therefore performed to identify studies which had applied economic techniques to health care delivery, with a focus on patients' experiences of care. Again, their searches were not restricted to those compatible with the QALY; they included specific search terms

for the methodologies (e.g. TTO, CA) within their search strategy. Their scope was therefore wider than the inclusion/exclusion criteria applied within this study.

Eighty-nine studies met their inclusion criteria. They found that studies had frequently estimated the value of the features of health care, such as location of treatment, travelling time, cost of travelling, waiting time, privacy. However, they also found that the implication of the features for patients, for example, how it makes them feel, was considered less often.

Following additional qualitative work, which is not related to the aims and objectives of this review, Ryan et al (2014) conclude that it is important to consider patients experiences of health care, and as shown by their exercises to value process attributes using standard economic measures such as the SG and VAS it is feasible to do so. They also concluded that the best approach for undertaking this is not yet known and further exploratory work is required.

The final review was by Higgins et al. (2014) who undertook a review to ascertain whether convenience matters in health care delivery. They accept the assumption that if only the health outcomes of an intervention are considered, there may be additional and important non-health attributes which are neglected, and this could potentially lead to a suboptimal provision of health care. Their review focused

specifically on studies reporting convenience aspects of process utility, for example preferences for waiting times or distance from care, and mode or frequency of medication administration. They included studies which could be used either in cost-utility analyses (TTO or SG), or cost-benefit analyses (WTP or contingent valuation). They identified 27 studies which met the inclusion criteria. They divided the studies into 2 categories: those relating to the administration of an intervention; those relating to ease of access to an intervention. Although some studies did use QALY compatible approaches, over half (56%) used DCE with a cost-related attribute or WTP valuations. SG was used in 2 studies (7%), and TTO was used in 3 studies (11%). They also found that the terminology used to describe non-health attributes differed considerably between studies, as did the estimates of process valued. However, all but one of their identified studies identified some component of convenience related process utility.

4.2.4 Discussion

This literature review was undertaken to identify published studies which provide a measure of process utility which can be incorporated into the QALY calculation. The focus on QALY compatible studies was identified as particularly important due to the wide-spread use of the QALY, driven predominantly by the requirements

dictated by regulatory bodies such as NICE. The identification of studies through conventional methods was difficult due to the lack of consistent indexing and terminology across the studies. This resulted in several being identified through ad hoc methods, and it is therefore acknowledged that there may be additional studies which the searches failed to identify.

Given the focus of this thesis the literature review aimed to concentrate only on those studies reporting an "empirical estimate of process utility which could be incorporated into the QALY." This introduced challenges in setting the inclusion and exclusion criteria. For example, standard conjoint analyses were excluded; however, chained conjoint analyses which did report a utility estimate were included. Although the study aimed to include only studies which reported a utility estimate anchored on the scale where 0 represents dead, and 1 represents full health, not all of the studies did use these endpoints. For example, Birch (2003) used an upper anchor of full health, but a lower anchor of invasive cancer requiring hysterectomy; In comparison Boye (2011) used an upper anchor of perfect health, and a lower anchor of "worst health represented by a person with T2DM, blood glucose levels not in control, body weight 10% higher than current weight and suffering from multiple complications." Therefore despite the aim of identifying studies reporting comparable utility estimates, there were evident differences in

the approaches taken. However, it is likely the results are still more focused and relevant to the aims of this thesis, than they would have been if these inclusion and exclusion criteria had not been applied.

Those studies which were identified demonstrated that whilst there is no universally accepted methodology for the detection and measurement of process utility, evidence does support its existence. However, when drawing conclusions from the review it is important to consider the extent of publication bias, where studies resulting in negative findings are less likely to be published or to receive pharmaceutical funding to be pursued to publication.

Utility Estimates

The review identified studies associated with processes involved in the treatment of patients (through both drug administration and surgical procedures), preventative care (prevention of hip fractures) and the screening of individuals for the prevention and early detection of morbidities. The utility estimates showed considerable variation, with estimates for treatment processes ranging from 0.001 for both comparisons of a laparoscopic cholecystectomy with an EWSL procedure in gallstone disease (Cook et al. 1994) and as a decrement for medications which have to be taken on an empty stomach (Kauf et al. 2008), to 0.09 for both inhaled versus

injectable medication in diabetes (Chancellor et al. 2008), and injections every 2 weeks versus every 3 months in haematology treatment. There was also variation between measurements of the same processes identified within different studies. For example, both Kauf et al. (2008) and Boye et al. (2011) compared patient preferences for medication which needs to be taken with food or on an empty stomach, and medication which can be taken irrespective of the meal times. Boye et al. (2011) estimated a utility increment of 0.006 for medication which can be taken at any time of the day, this compares to Kauf et al's (2008) decrements of 0.0028 for medication which must be taken with food, and 0.001 for medication which must be taken on an empty stomach.

Due to the range of methodological approaches used to identify and measure process utility it is difficult to know whether these differences in estimates are a true reflection of the amount of process utility which enter into an individual's utility function, or whether they are associated with features of the studies' methodological design. Without a standardised approach to the methodology for the detection and measurement of process utility comparisons between estimates are difficult and should be viewed with caution.

The statistical significance of estimates should also be considered. Not all estimates reached levels of statistical significance, for example the estimates reported by Kauf et al. (2008) for the different treatment regimens in patients receiving HIV treatment. In addition to this many are below the minimally important difference in utilities suggested by Walters and Brazier (2003) which ranges from 0.010 to 0.048. Other estimates seem large when compared to existing measures of utility for certain conditions. For example, Osbourne et al. (2007) report a process utility value of 0.27 for a move away from subcutaneous administration of haematology treatment to oral administration. This is a similar magnitude to the difference between perfect health, and for example anorexia, a severe hearing disorder or permanent cognitive impairment after bacterial meningitis reported by Stothard et al. (1997), who report utility estimates between 0.7 and 0.8. These 2 points were highlighted by one of the referees of the Brennan & Dixon (2013) paper.

Study populations.

The populations used to value health state utilities varied. Fourteen of the 20 included studies used patients, 6 used a general population sample, and one used a general population sample, with sub analyses on patients with either T1DM or T2DM (Evans et al. 2013). Previous research indicates that utilities assigned to the

same health states vary substantially according to the group of individuals valuing them, with patients typically producing higher values than the general population (Boyd et al. 1990). This added heterogeneity between studies introduces additional uncertainty to the process utility decrements identified.

Methodology

Hypothetical Health States

Seventeen of the 20 included studies developed hypothetical health states which maintained constant health outcomes and differed only in descriptions of process. These health states were subsequently valued using one of the various approaches (SG, TTO, WTO) to provide preference measures for pairs of health states with and without process for comparison. Brazier et al. (2007) suggest that although a lot of research has been undertaken into the methods for valuing health states, less emphasis has been placed on methodological issues associated with describing health, despite the validity of the description of the health state impacting on the validity of the utility estimates derived. Some authors reported the methods used to develop the health state descriptions, for example, Boye et al. (2011) drafted type 2 diabetes related health states which a physician then reviewed. The revised health states were then piloted on 11 patients in a valuation questionnaire. This

lead to further revisions, and a second phase of piloting was then performed to ensure no further amendments were required before launching the study. Other authors provide no detail on how the health state descriptions were developed or validated (Birch et al. 2003; Chancellor et al. 2008).

Although this approach allows a clear comparison of health states with and without process, it does have its limitations. Brazier et al. (2007) suggest that at times there can be poor linkage between health state descriptions and clinical evidence. It is difficult to capture the potential uncertainties surrounding included components (symptoms, functioning, wellbeing) within a small number of health state descriptions. They suggest that the validity of health states and scenarios needs to be rigorously assessed (Brazier et al. 2007). A further limitation is the presence and impact of focusing effects. This occurs when the respondents valuing the health states place too much emphasis on the component being studied, for example process utility. This can lead to an over emphasis on the valuation of that one component, and impact on the valuation of the health state as a whole (Brazier and Rowen, 2011).

The use of bespoke hypothetical health states contradicts both the NICE Methods guidance (2013) and recommendations from the Washington Panel (Siegel et al.

1997), both of which report a preference for the use of health states based on validated health related quality of life measures. Estimates of process utility developed based on valuations of hypothetical health states developed using expert or patient opinion are therefore incompatible with NICE and Washington Panel reference case methods.

Population

Both the NICE Methods Guide (2013) and the Washington Panel (Siegel et al. 1997) also have a strong theoretical preference for values to be based on those of the general population. However, the majority of the studies we identified, and of those identified by Opmeer et al. (2010) used patient respondents. It could be argued that only patients could fully appreciate the specialised nature of certain illnesses or treatments, however, studies have identified systematic differences between patient and general population values (Brazier et al. 2005). Therefore the resulting estimates of process utility obtained through patient valuations are incompatible with Washington Panel and NICE reference case methods.

Valuation method

Several of the included studies (Boye et al. 2011; Cook et al. 1994; Chancellor et al, 2008; Birch et al. 2003; Matza et al. 2013; Howard et al. 2003) used a two stage or chaining methodology for the valuation of health states, where intermediate health states are measured relative to the best health state and the worst state. This methodology put forward by Torrance (1986) to overcome issues associated with the valuation of temporary health states, leads onto an additional point for consideration: the issue of how the process utility estimates are applied within cost-effectiveness analyses (Gafni and Zylak 1990). In order to apply the utility estimates for example in an economic model, assumptions must be made on whether the effects are deemed to be persistent (and so extend beyond the actual process), or whether they are transient (and so apply only to the time during which the process occurs). An example of this may be a cost-effectiveness analysis into the use of inhaled insulin versus injectable insulin. Applying the utility gain only during the time which insulin is administered versus its application to the associated health state will have a considerable effect on the estimated QALY gains from the improved process. If the former approach is taken then the QALY gains for some processes may be very small. The study by Johnson et al. (1996) for example, identifies a utility decrement of 0.223 associated with moving from IV to oral drug

administration, which seems implausible for an ongoing reduction in utility, but possible for a utility reduction on the day of treatment. It is therefore important that studies report the timeframe used within many of the trade-off questions. This allows the utility estimate to be adjusted for duration before being applied correctly within the QALY calculation.

Future Research

Whilst the majority of studies used recognised TTO and SG techniques for the valuation of health states, one study used generic preference based measures (PBM) to determine estimates of process: Kauf et al. (2008) used responses from the SF-36. This raises the question, to what extent is process utility captured within existing PBMs? If process is already captured then adding further allowances to QALY estimates will lead to a double counting of the effects. This therefore is an important issue which needs further research.

More rigorous testing of the validity of the methods used is also essential in future studies. This could take the form of psychometric tests within the stated preference studies themselves and comparison with non-preference based patient reported outcome measures from patients. Interestingly, the study of inhaled insulin by Chancellor was 'tested' through the launch of the related product –

Exubera – and observed patient uptake (Chancellor et al. 2008). Whilst the stated preference study suggested important utility gains from moving to inhaled insulin, patients were very reluctant to use it and the product was eventually withdrawn from the market. Therefore patients' revealed preferences suggested a utility decrement associated with using inhaled insulin.

This review has introduced some of the limitations to using hypothetical health states in utility estimation. These include uncertainty surrounding the validity of the health state descriptions, problems linking descriptions to clinical evidence, and potential framing effects. One approach to overcome the difficulties of using bespoke health state descriptions is to bolt-on dimensions to existing preference based measures (Brazier and Rowen, 2011). This approach has been considered for the addition of health domains (cognition, sleep and pain), (Krabbe et al. 1999; Yang et al. 2011; Brazier et al. 2010), but it would also be possible to add on bespoke process descriptions, such as those included in the scenarios developed by Ryan et al. (2014), or descriptions from a validated measure of patient experience such as the PEQ (Stein et al. 2001) which was used in the ePAQ EEACT reported in Chapter 3.

4.2.5 Conclusion

The review reported within this chapter aimed to identify published studies which provide measures of process utility using methodologies which allow resulting estimates to be incorporated into QALY calculations. The studies identified demonstrate that the process of health care has been considered in relation to what enters an individual's utility function, and the evidence to date does support its existence, not only in the administration of drugs, but also in approaches to screening and preventative care. This review indicates that patients place value on certain aspects of the process of receiving health care.

When considering the number of economic evaluations that are published, in relation to the number identified from this search, it can be assumed that this aspect of an individual's utility function is rarely included in cost-utility analyses, and therefore many of these studies may be presenting results that do not truly reflect the impact on patient's health-related quality of life, and so are not an accurate estimation of cost-effectiveness.

Health economists therefore need to acknowledge that patients attach value to process aspects and other non-health outcomes of health care interventions. As there is no established approach to incorporating these additional outcomes into

the QALY, and one approach may be to undertake alternative studies using for example discrete choice experiments, and use the findings in combination with the QALY outcomes. However, recognising the importance of process and non-process outcomes also recognises the need to develop these QALY-friendly approaches further.

This review has highlighted areas for further research, including a comparative study of alternative methods for the measurement of process utility, the need for testing the validity of results through psychometric approaches, and comparative studies with other patient reported measures, and the bolting-on of a process domain to an existing generic preference based measure (E.g. EQ-5D or SF-6D) using an already validated measure of patient experience. Chapters 5 and 6 explore this further.

Health economists and decision makers therefore need to acknowledge that aspects of the process of receiving health care may enter into a patient's utility function and consideration should be given on a study by study basis as to whether this should be incorporated into a cost-utility analysis.

4.3 References

- Attema, AE. Edelaar-Peeters, P. Versteegh, MM. Stolk EA. (2013). Time Trade-off: one methodology, different methods. *European Journal Health Economics* 14(Suppl 1), S53-S64.

- Birch, S. Melnikow, J. Kuppermann, M. (2003) Conservative versus aggressive follow up of mildly abnormal Pap smears: testing for process utility. *Health Econ.* 12(10), 879-84.

- Boye, KS. Matza, LS. Walter, KN. Van Brunt, K. Palsgrove, AC. Tynan, A. (2011) Utilities and disutilities for attributes of injectable treatments for type 2 diabetes. *Eur J Health Econ*. 12(3), 219-30.

- Brazier, J. Akehurst, R. Brennan, A. Dolan, P. Claxton, K. McCabe, C. Sculpher, M. Tsuchyia, A. (2005) Should patients have a greater role in valuing health states? *Appl Health Econ Health Policy*. 4(4), 201-8.

- Brazier, J. Rowen, D. Tsuchiya, A. Yang, Y. Young, T. (2010) What a pain: adding a generic dimension to a condition-specific preference-based measure. *HESG Abstract*.

- Brazier, JE. Rowen, D. NICE DSU Technical Support Document 11: Alternatives to EQ-5D for generating health state utility values. Available from: http://www.nicedsu.org.uk [Accessed 16th September 2015].

- Brennan, VK. and Dixon, S. (2013) Incorporating Process Utility into Quality Adjusted Life Years: A Systematic Review of Empirical Studies, Pharmacoeconomics, 31(8):677-91

- Brouwer, WB. van Exel, NJ. van den Berg, B. van den Bos, GA. Koopmanschap, MA. (2005) Process utility from providing informal care: the benefit of caring. *Health Policy*. 74(1):85-99.

- Cairns, J. Shackley, P. Hundley, V. (1996) Decision Making with Respect to Diagnostic Testing: A method of Valuing the Benefits of Antenatal Screening. *Med Decis Making,* 16, 161.

- Chancellor, J. Aballéa, S. Lawrence, A. Sheldon, R. Cure, S. Plun-Favreau, J. Marchant, N. (2008) Preferences of patients with diabetes mellitus for inhaled versus injectable insulin regimens. *Pharmacoeconomics.* 26(3), 217-34.

- Cook, J. Richardson, J. Street, A. (1994) A Cost utility analysis of treatment options for gallstone disease: methodological issues and results. *Health Economics*, 3, 157-168.

- Donabedian, A. (1988) Evaluating the quality of medical care. Millbank Fund Quarterly: Health and Society, 44. Cited in McAlister D. (1994) Putting health economics into quality. *Public Money & Management*. 14(2), 15-22.

- Donaldson, C. Birch, S. (2014) Inclusiveness and health economics. *Soc Sci Med*, 108, 246-247.

- Donaldson, C. Shackley, P. (1997) Does "process utility" exist? A case study of willingness to pay for laproscopic cholcystectomy. *Soc Sci Med*, 44, 699-707

- Entwistle, V. Firnigl, D. Ryan, M. Francis, J. Kinghorn, P. (2012) Which experiences of health care delivery matter to service users and why? A critical interpretative synthesis and conceptual map. *Journal of Health Services Research and Policy,* 17(2), 70-78.

- Evans, M. Holm Jensen, HH. Bogelund, M. Gundaard, J. Chubb, B. Khunti,, K. (2013) Flexible insulin dosing improves health-related quality-of-life (HRQoL): a time trade off survey. *Journal of Medical Economics*, 16(11), 90.

- Gafni A, Zylak, CJ. (1990) Ionic versus non-ionic contrast media: a burden or a bargain? *Can Med Assoc J,* 143(6),475-478.

- Gerard, K. Mooney, G. (1993) QALY league tables: handle with care. *Health Economics,* 2(1), 59–64

- Gonzalez Block, MA. Sandiford, P. Ruiz, JA. Rovira J. (2001) Beyond health gain: the range of health system benefits expressed by social groups in Mexico and South America. *Social Science and Medicine*, 52,1537-1550.

- Higgins, A. Barnett, J. Meads, C. Singh, J. Longworth, L. (2014) Does Convenience Matter in Health Care Delivery? A Systematic Review of Convenience-Based Aspects of Process Utility. *Value in Health*, 17,877-887.

- Howard, K. Salkeld, G. McCaffery, K. Irwig, L. (2008) HPV triage testing or repeat Pap smear for the management of atypical squamous cells (ASCUS) on Pap smear: is there evidence of process utility? *Health Econ*, 17(5), 593-605.

- Hutchins, R. Viera, AJ. Sheridan, SJ. Pignone, MP. (2015) Quantifying the utility of taking pills for cardiovascular prevention. *Circ Cardiovasc,* 8,155-163.

- Johnson, ES. Sullivan, SD. Mozaffari, E. et al (1996) A utility assessment of oral and intravenous ganciclovir for the maintenance treatment of AIDS-related cytomegalovirus retinitis. *Pharmacoeconomics*, 10, 623–9.

- Kauf, TL. Roskell, N. Shearer, A. Gazzard, B. Mauskopf, J. Davis, EA. Nimsch, C. (2008) A predictive model of health state utilities for HIV patients in the modern era of highly active antiretroviral therapy. *Value Health,* 11(7),1144-53.

- Krabbe, P. Stouthard, M. Essink-Bot, M. Bonsel, G. (1999) The effect of adding a cognitive dimension to the EuroQol Multiattribute Health-Status Classification System. *J clin Epidemiol,* 52(4), 293-301.

- Matza, LS. Cong, Z. Chung, K. Stopeck, A. Tonkin, K. Brown, J. Braun, A. Brunt, KV. McDaniel, K. (2013) Utilities associated with subcutaneous injections and intravenous infusions for treatment of patients with bone metastases. *Patient preference and Adherence*, 7, 855-865.

- McAlister D. (1994) Putting health economics into quality. Public Money & Management, 14(2),15-22

- Mcguire, A. Hendersen, J. Mooney, G. (1998) The Economics of Health Care: An Introductory Text. London, Routledge.

- Mooney G. (1998) Beyond health outcomes: The benefits of health care. *Health Care Analysis*, 6(2), 99-1.

- NICE (2013) National Institute for Health and Care Excellence. Process and Methods Guides: Guides to the methods of technology appraisal 2013. Available from: http://publications.nice.org.uk/pmg9 [Accessed 28th August 2015].

- Oliver, A. (2004) Prioritizing Health Care: Is "Health" Always an Appropriate Maximand? *Med Decis Making*, 24, 272.
- Opmeer, BC. de Borgie, CA Mol, BW. Bossuyt, PM. (2010) Assessing Preferences Regarding Healthcare Interventions that Involve Non-Health Outcomes: An Overview of Clinical Studies. *Patient*, 3(1), 1-10.

- Osborne, RH. De Abreu Lourenço, R. Dalton, A. Houltram, J. Dowton, D. Joshua, DE. Lindeman, R. Ho, PJ. (2007) Quality of Life Related to Oral versus Subcutaneous Iron Chelation: A Time Trade-off Study. *Value Health*, 10(6), 451-6.

- Osbourne, RH. Dalton, A. Hertel, J. Schrover, R. Smith, DK. (2012) Health-related quality of life advantage of long-acting injectable antipsychotic treatment for schizophrenia: a time-trade-off study. *Health and Quality of Life Outcomes,* 10, 35.

- Polster, M. Zanutto, E. McDonald, S. Conner, C. Hammer, M. (2010) A comparison of preferences for two GLP-1 products-liraglutide and exenatide-for the treatment of type 2 diabetes. *J Med Econ*, 13(4), 655-61.

- Ryan, M. Kinghorn, P. Entwistle, VA. Francis, JJ. (2014) Valuing patients' experiences of healthcare processes: Towards broader applications of existing methods. *Social Science and Medicine* 106,194-203.

- Sadigh, G. Carlos, RC. Kazerooni, EA. Kelly, AM. (2013) Patient preferences for coronary computed tomography angiography versus conventional catheter angiography for the diagnosis of coronary artery disease. *Academic Radiology,* 20(9)1091-1098.

- Salkeld, G. (1998) What are the benefits of preventative care? *Health Care Analysis,* 6,106-112.

- Salkeld, G. Quine, S. Cameron, ID. (2004) What constitutes success in preventive health care? A case study in assessing the benefits of hip protectors. *Soc Sci Med.* 59(8),1593-601.

- Schmier, JK. Palmer, CS. Flood, EM. Gourlay, G. (2002) Utility assessments of opioid treatment for chronic pain. *Pain Med*, 3(3):218-30.

- SiegeI, JE. Torrance, GW. Russell, Luce, BR. Weinstein MC. Gold, MR. Members of the Panel on Cost Effectiveness in Health and Medicine. (1997) Guidelines for Pharmacoeconomic Studies. Recommendations from the Panel on Cost Effectiveness in Health and Medicine. *Pharmacoeconomics,* 11(2), 159-168.

- Steine, S. Finset, A. Laerum, E. (2001) A new, brief questionnaire (PEQ) developed in primary health care for measuring patients' experience of interaction, emotion and consultation outcome. *Family Practice,* 18,410-418.

- Stouthard, ME. Essink-Bot, ML. Bonsel, GJ. Barendregt, JJ. Kramer, PG. van de Water, HP. Gunning-Schepers, LJ. van der Maas, PJ. (1997) Disability weights for diseases in the Netherlands. Rotterdam: Erasmus University.

- Swan, J. Sainfort, F. (2003) Process Utility for Imaging in cerebrovascular disease. *Acad Radiol,* 10(3), 266-74.

- Swan, JS. Fryback, DG. Lawrence, WF. Sainfort, F. Hagenauer, ME. Heisey, DM. (2000) A Time-trade-off method for cost effectiveness models applied to radiology. *Med Decis Making*, 20:79.

- Swan, JS. Lawrence, WF. Roy, J. (2006) Process utility in breast biopsy. *Med Decis Making*, 26(4),347-59.

- Tavakoli, M. Davies, H. Thompson, R. (2000) Decision analysis in evidence-based decision making. *Journal of Evaluation in Clinical Practice*, 6(2), 111-120.

- Torrance, GW. (1986) Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5,1-30.

- Walters, SJ. Brazier, JE. (2003) What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. *Health Qual Life Outcomes*, 1,4.

- Yang, Y. Brazier, J. Tsuchiya, A. (2008) The effect of adding a 'sleep' dimension to EQ-5D. *Health Economists' Group meeting,* January.

Chapter 5: Summary of bolt-on studies and the development of a methodology

5.1 Introduction

This chapter explores an approach for expanding on the traditional QALY, through the bolting on of a domain to an existing generic measure of quality of life. The decision to adapt the QALY, as opposed to the utilisation of alternative direct approaches e.g. willingness to pay, was taken for two main reasons. Firstly, as previously stated, the use of the QALY is the dominant approach to economic evaluation in England, and is widely used and understood by economists, clinicians and commissioners. Therefore, expanding our knowledge and application of the QALY was felt to be the most applicable to current practise. And secondly, when reflecting on the ePAQ trial, the fact that a generic preference based measure (SF-12) was included; that there were concerns that this failed to capture important differences in process; but that this information was captured within additional questionnaires (PEQ, CSQ) lends itself to a bolt-on approach, where already collected data can be utilised to create the bolt-on, and to provide estimates of cost-effectiveness pre and post the incorporation of process.

It is also acknowledged that the SF-12 is not the "preferred" measure of health related quality of life in adults by NICE (NICE, 2013, p32). However, NICE also accept non-preference based methods, with suitable justification for the approach taken, which in this case was the reduced sensitivity at the top end of the EQ-5D scale, when compared to the SF-12, combined with a study population of relatively young and mobile women. This shows that the EQ-5D is not always the preference based measure of choice, and therefore research should not be solely focused on the EQ-5D.

This chapter initially provides an overview of currently published bolt-on studies, and then applies this to both the concept of process utility, and the ePAQ EEACT reported in Chapter 3. A methodology is then developed, which aims to capture the process utility associated with the ePAQ trial arms, and ultimately will result in updated cost effectiveness outcomes which incorporate process. The methodology is reported in this chapter, and the study results are reported in Chapter 6.

5.2 Background

The NICE induced increase in economic evaluations of health care over the last decade has led to an increase in the use of generic measures of quality of life (e.g. EQ-5D, SF-36). Greater utilisation of these generic measures has highlighted that

for some diseases they fail to capture important dimensions of health. As a result of this, several health economists have attempted to add dimensions to existing instruments. This section provides a summary of the current literature on bolt-on studies performed in health economics. All studies reporting a bolt-on study performed within health economics were included. They were identified through a literature search performed in Pubmed, and through contacting experts in the field to check whether they were aware of additional papers (John Brazier and Yaling Yang). As several of the studies were performed at ScHARR I was confident that the key papers were identified using this strategy. The initial searches were performed in October 2014, and updated in July 2015.

Studies have been performed in 3 main treatment areas: cognition; sleep; pain. Summary details of these studies are reported in Table 13, and these studies will now be discussed in greater depth.

5.2.1 Summary of "bolt-on" studies.

Krabbe et al. (1999)

Krabbe et al (1999) undertook a study to add a cognitive dimension to the EQ-5D-3L. This is the 3 level version of the measure which includes 5 domains, each with 3 levels: No problems; some problems; severe problems. The EQ-5D is commonly used as the source of utility estimates in economic evaluations, particularly as it is the preferred measure of health related quality of life by NICE (NICE, 2013). However, certain aspects of the EQ-5D have been questioned, specifically in relation to the use of the tool in dementia treatment. Krabbe et al. (1999) discuss the relevance of the EQ-5D for use in economic evaluations in dementia considering the absence of a cognition domain. They therefore undertook a study to bolt-on a cognitive domain to the EQ-5D. They assumed that adding a cognitive dimension would increase the comprehensiveness of the tool (to include a wider definition of health status), but this may be at the cost of increased complexity to the user and therefore greater variability in responses.

Krabbe et al. (1999) maintained the standard EQ-5D-3L classification system, and added to this a 6th cognition domain also with the 3 levels (no impairment of cognitive functioning, some impairment of cognitive functioning and severe

impairment of cognitive functioning). This modified classification system allowed 729 health state descriptions. They also added components to the "usual activities" domain to cover cognitive activities such as "memory, concentration, coherence and IQ" (Krabbe et al. 1999, p297). Fourteen health states were selected, ensuring that they included an even mix of good, moderate, and bad health states. The sixth cognitive attribute was added to all 14. The best (11111(1)) and worst (33333(3)) health states were also included. Two additional EQ-5D states were added to construct a parallel set of EQ-5D+C (EQ-5D + Cognition) health states with an assumed significant effect of the cognitive functioning attribute. Two additional health states were included which differed only in the cognition level (by 1 level) from health states already included in the study. This allowed the effect of varying the cognition domain by 1 level to be explored with differing levels of background severity. The resulting study therefore contained:

- 18 health states for the EQ-5D
- 20 health states for the EQ-5D+C

Health state descriptions were valued using the standard VAS ranging from 0 to 100, where 0 is the "worst imaginable health state" and 100 is the "best imaginable health state". Respondents were asked to locate the health state descriptions on

the scale between the two anchors so that the intervals between the scores corresponded to the differences they perceive.

The distribution of EQ-5D questionnaires and EQ-5D+C questionnaires was performed randomly. Analysis of the data included:

- Calculations of means and standard deviations for all valued health-state descriptions;
- Statistical testing within pairs (EQ-5D versus EQ-5D+C);
- Paired t-tests were performed for some health states i.e. replication health states; health states with level variations.

Results of *t* tests for differences between the EQ-5D and EQ-5D+C showed significant differences in 2 cases: health state 11211(2) where a “moderate” cognition level was added to a “good“ state, and health state 12311(3) where a “bad“ cognition level was added to a “moderate“ state. Adding a “good” cognition level to a “bad” health state did not result in any meaningful improvements. Adding a “bad” cognition level to a “bad” health state resulted in a poor valuation. Therefore indicating that “good” health states were affected more by the addition of impaired cognition, but the reverse was not true for adding a “bad” health state.

The authors reported that “generally, the standard deviations for valuations elicited by the EQ-5D+C version were slightly greater than those for the EQ-5D.”(p297).

Authors included some health states twice in the study [11111(1), 33333(3)] and compared valuations for both pairs within each version as a reliability test. Paired *t* tests showed no difference in valuations of 11111(1) for either version but a small but significant difference in valuation for health state 33333(3) for the EQ-5D+C version. Additional checks were made where they included 2 health states twice changing only one level. The valuations of these health states were as expected.

The authors concluded that the approach they developed to add the cognition domain to the existing EQ-5D was suitable for judging the revised EQ-5D+C tool, and that those who are developing multi-attribute health status tools in the future should be aware of the omission of cognition in the existing tools available.

Wolfs et al. (2007)

Wolfs et al. (2007) used the EQ-5D+C developed by Krabbe et al. (1999) and the EQ-5D in a population of patients with cognitive impairment, and compared the validity of both tools through assessment of construct validity. This refers to the extent to which a measure correlates to other relevant measures. They compared the EQ-5D+C to the MMSE (mini-mental state examination). Patients included in

their study were those who participated in a larger RCT, the MEDICIE study, which compared the effects of a multi-disciplinary diagnostic observation centre for psycho-geriatric patients with usual care (GP review). Within the trial, the MMSE, EQ-5D and VAS were administered to patient proxies, who were then asked to value the additional cognitive domain, and complete the VAS. Patient proxies also completed the MMSE. They hypothesised that the correlation between changes in the MMSE and EQ-5D+C would be greater than the correlation between the changes in the MMSE and EQ-5D.

The study found that the correlations between the EQ-5D and the MMSE were similar to the correlations between the EQ-5D+C and MMSE. In addition to this, both the EQ-5D and EQ-5D+C were responsive to changes in MMSE, with the EQ-5D being slightly more sensitive. Their hypothesis was therefore rejected, and the study concluded that there was no need to expand on the EQ-5D to add the cognitive dimension.

The authors did however, discuss the limitations to their study, namely, that the exploration into the two tools was performed within a RCT, which was not designed purely for this purpose, and therefore improvements could have been made to the methodology, for example, in order to reduce the burden on trial participants (and

patient proxies), they were only asked to complete the EQ-5D, and then the additional cognition domain. They were not asked to complete the EQ-5D+C as a whole. They also suggest this approach may have led to respondents focusing on the cognition domain, and may therefore have introduced framing effects. The use of proxies was discussed, although they concluded that previous research has indicated that the use of a proxy in assessing health related quality of life is a valid alternative to patients with dementia. And finally they discussed their measures for comparison. The comparisons of the EQ-5D and EQ-5D+C were not based on utilities, as it is not possible to obtain comparable utilities from the EQ-5D+C measure. The utilities which do exist are based on valuations by health professionals as opposed to the general population, and were obtained using PTO.

Yang et al. (2008)

Yang et al. (2008) discuss the possible impacts of poor sleep on people's production and performance in society, and on their quality of life. They therefore undertook a study examining the effect of adding a sleep dimension to the EQ-5D. The additional dimension maintained the 3 levels within the EQ-5D-3L resulting in:

- I have no problems with sleep;
- I have some problems with sleep;
- I have extreme problems with sleep.

Eighteen EQ-5D + sleep states, and 16 EQ-5D states were selected for valuation, and following stratification according to severity groups, were randomly allocated to 2 blocks, each containing 8 states. The health states included 6 matched health state pairs; which contained an EQ-5D state and a corresponding EQ-5D+sleep state, where the only difference was the addition of the sleep domain.

The EQ-5D and EQ-5D + sleep health states were then valued by a general population sample of 160 people using TTO techniques. Comparisons between paired states suggested that adding the sleep dimension had no influence on valuations of the original EQ-5D health state.

Brazier et al. (2011)

Brazier (2011) undertook a similar study using a condition specific measure. They explored the impact of adding a pain dimension to the AQL-5D. The AQL-5D is an asthma specific preference based measure derived from the Asthma Quality of Life Questionnaire (AQLQ). The original AQLQ contains 5 domains (concern about asthma, shortness of breath, weather and pollution stimuli, sleep impact and activity limitations), and each have 5 levels (1 = no problem to 5 = severe problems). When mapped onto the AQLD-5D this resulted in a health state classification system with 3 level of severity (1 = no problems; 2 = some problems; 3

= extreme problems), this reduced health-state classification system was used in the study. Brazier et al. (2011) added the pain domain from the EQ-5D to the classification system to create the AQL-6D. They then performed a valuation study on both the AQL-5D and the AQL-6D. Sixteen health states were selected from the AQL-5D and 18 health states were selected from the AQL-6D. These included 4 health states that were matched across each measure. Members of the public valued 8 health states from either the AQL-5D or AQL-6D using TTO. The impact of adding the dimension was then examined by comparing the mean values for the matched states (independent samples t tests), and by modelling the data and comparing co-efficients across the AQL-5D and AQL-6D.

The results showed that the additional pain dimension had a significant coefficient, and that it impacted significantly on the coefficients of the other dimensions, with the degree of the impact differing by dimension and severity level. Brazier et al. (2011) applied their findings to clinical trial data in a trial which had used both the AQLQ and the EQ-5D. Values generated for the AQL-5D and AQL-6D were compared in terms of mean scores for all patients included in the trial, and also for sub-samples of patients identified according to asthma symptom scores. The results showed that the mean health state value produced from the ALQ-5D was consistently lower than for the AQL-6D value across the 5 asthma symptom severity

groups. For the pain dimension, the AQL-6D group mean scores only exceeded the AQL-5D scores for extreme pain.

The results therefore indicated that overall, the addition of a pain/discomfort domain to a health state resulted in lower utility estimates. This is opposite to the expected result. Brazier et al. (2011) put forward suggestions for this, they suggested on valuing the AQ-5D with no pain mentioned, respondents may assume there is associated pain when an equivalent AQ-6D health state reporting no pain is valued. They also suggest that a simplifying heuristic is used, whereby respondents focus on one dominant dimension. Brazier et al. (2011) suggest that respondents should be asked additional questions in regard to these observations to understand these issues more fully, as they accept that they all raise serious concerns about the health state valuations.

Brazier et al. (2011) also found that the addition of a pain domain at level 1 (no problems) significantly increased the mean health state value. This highlights an assumption that should be made that "no problems" as Level 1 in a bolt-on does not alter the mean value given to any health state that includes Level 1 in the bolt-on. This is because the additional information should be irrelevant to the valuation task. This is akin to the assumption of independence of irrelevant alternatives (IIA)

that is used in decision theory more generally (Alava et al. (2013). Brazier's et al's (2011) findings show that this assumption doesn't hold, and suggests that the assumptions about rational choice are flawed, potentially through the use of the simplifying heuristics discussed above. Swinburn et al. (2013)

A similar study was undertaken by Swinburn et al. (2013), who explored adding a disease-specific bolt-on to the ED-5D-5L related to psoriasis. They termed this the EQ-PSO questionnaire. Their study consisted of 2 components: the development and psychometric testing of the bolt-on instrument, and the subsequent valuation study.

The bolt-on instrument was developed using a literature review and clinician and patient interviews. Initially 4 new dimensions were developed and examined through patient interviews to ascertain whether the concepts were important to patients. Psychometric testing of the dimensions was then undertaken. This resulted in 2 of the 4 dimensions being selected (skin irritation and self-confidence). During the second stage, the EQ-PSO questionnaire bolt-on instrument was valued by a convenience sample in the UK using TTO and VAS. Their study found that the new questionnaire was better at predicting psoriasis outcomes than some existing disease specific measures.

Swinburn et al. (2013) suggest that bolt-on's should only be considered where the existing measures have shown weakness. This is certainly the case for the use of the SF-6D in the ePAQ EEACT. They also suggest further qualitative work to ensure the information which is important to patients is captured.

Yang et al. (2015)

A final study explored the impact of 3 bolt-ons' to the EQ-5D. These were developed for vision problems, hearing problems and tiredness. Each dimension had 3 severity levels, in line with the EQ-5D-3L. Health states with and without the bolt-on dimensions were valued by 300 members of the UK general public using TTO methodology. Mean health state valuations were then compared, and regression was used to explore the impact of the different dimensions and underlying health state severities.

The study found that the bolt-on dimensions did have an impact on the health state valuations. The degree of the impact was dependent on the severity of the underlying EQ-5D health state, and the bolt-on dimension levels. They therefore suggest that a decrement alone may not be enough to reflect the relationship between the added domain and existing dimensions, and an additive model may not be sufficient.

Table 13. Summary of bolt-on studies.

Author (date)	Existing tool	Added Domain	Methodology	Valuation Method	Tariff	Results	Advantages	Disadvantages
Brazier (2011)	AQL-5D – asthma specific measure	Pain AQL-6D	Valuation study and regression analysis AQL-5D and AQL-6D were applied to clinical trial data	TTO	No	▪ Addition of pain impacted significantly on co-efficients of dimensions, but differed by dimension and severity level. ▪ When relating back to the trial, the mean HS value produced from the AQL-5D was consistently lower than for the AQL-6D across the 5 asthma symptom severity groups.	▪ Use of patient data with both measures (EQ-5D and AQL-5D)	▪ Potential focusing effect ▪ Explored only an additive model ▪ Results not generalisable
Krabbe (1999)	EQ-5D	Cognition EQ-5D+C	Comparison of paired health states, valued using VAS.	VAS	No	▪ 87 questionnaires returned. ▪ Cognitive dimension generated systematically different values compared with the standard EQ-5D.	▪ Maintains underlying 3 level system of the EQ-5D.	▪ Valuation population university faculty members (they may place more emphasis on cognition).Simple additive model inadequate, due to dependencies between

Author (date)	Existing tool	Added Domain	Methodology	Valuation Method	Tariff	Results	Advantages	Disadvantages
						▪ Good states affected more, bad states less affected.		specific levels of different attributes.
Swinburn et al., (2013)	EQ-5D	2 domains: Skin irritation and self-confidence EQ-PSO	Literature review. Clinician and patient interviews. Psychometric testing. Valuation study.	TTO	No	▪ The revised EQ-PSO questionnaire captured more variance than the EQ-5D. ▪ It was also better at predicting psoriasis outcomes than some existing disease specific measures.	▪ Maintains underlying 5 level system of the EQ-5D. ▪ Has undergone psychometric validation ▪ Paper included a valuation study using TTO.	▪ The literature review used to develop the domains was not systematic. ▪ Domains were validated by only 8 patient interviews and one clinical specialist. ▪ Unable to translate to utilities due to differences in valuation studies for EQ-5D-5L and valuation study performed.
Wolfs	EQ-5D	Cognition	Application of	VAS	No	▪ 234 patients included. ▪ Construct validity –	▪ Application of Krabbe (1999) study in clinical practice	▪ Exploration of new tool within existing trial (not designed

Author (date)	Existing tool	Added Domain	Methodology	Valuation Method	Tariff	Results	Advantages	Disadvantages
(2007)		EQ-5D+C	Krabbe (1999) in practice			similar correlations found between EQ-5D and MMSE, and EQ-5D+C and MMSE. ▪ EQ-5D adequate to detect differences, little gained by using EQ-5D+C ▪ EQ-5D was slightly more responsive than the EQ-5D+C		specifically for this purpose) ▪ Completion by proxies. ▪ Focusing effect
Yang (2008)	EQ-5D	Sleep EQ-5D+sleep	Valuation study and regression analysis	TTO	No	▪ Co-efficient for sleep dimension non significant. ▪ Inclusion of sleep had no impact on valuation of other dimensions. ▪ No need to add sleep dimension	▪ Maintains underlying 3 level system of the EQ-5D.	▪ Small sample size
Yang	EQ-5D	3 domains:	Valuation study	TTO	No	▪ The additional do-	▪ Maintains underlying	▪ Interviews undertak-

Author (date)	Existing tool	Added Domain	Methodology	Valuation Method	Tariff	Results	Advantages	Disadvantages
(2015)		Vision Hearing tiredness				mains did impact on the health state valuation. ▪ Differences were associated with the levels of the domain added and the severity of the underlying health state.	3 level system of the EQ-5D ▪ Reviews more than one dimension, and allows comparisons to be made between each as based on same methodology.	en in 1 area of UK, therefore may not be generalisable to UK population. ▪ Dimensions added to EQ-5D-3L, not 5L version – which may be more sensitive to changes.

The remainder of this chapter describes the development of a process utility bolt-on study. It draws on previously performed studies, both those identified in the literature review of empirical studies of process utility, and the bolt-on studies just described.

In contrast to the existing bolt-on studies which add a health domain, this study aims to add on a process domain. The bolting-on of a domain to an existing generic preference based measure, follows recommendations made by NICE (2013) and The Washington Panel (Siegel et al. 1997), as opposed to developing bespoke health state descriptions. This ensures the health states are evidence based, and balanced, for example, avoiding the overemphasis on any particular characteristic. The study also aims to use components of existing and established instruments (SF-12, PEQ, CSQ-8), which have been validated using psychometric testing, thereby testing for reliability and validity. This negates the requirement to perform additional psychometric testing as was performed by Swinburn et al. (2010). As these 3 measures were used in the ePAQ trial reported in Chapter 3, it also allows the findings from the valuation study to be linked directly to existing trial data, and will provide estimates of process utility with increased internal validity (how do we know that the differences identified are a true representation of what actually happened), as the results will be applied to the ePAQ trial data to determine the "true" causes of the outcomes that were observed. This study therefore furthers the common thesis theme of improving the relevance of trial outcomes to enhance the cost-effectiveness outcomes.

5.3 Objectives

The objectives of this process utility valuation study were:

- To develop a methodology which utilises the SF-6D, PEQ and/or CSQ-8 data collected within the ePAQ trial, to investigate the presence, and magnitude of process utility.
- To ensure the methodology will allow incorporation of the resulting utility estimates into the QALY calculation.
- To apply the revised utility estimates, which include an element of process, back to the ePAQ EEACT results, and report an updated incremental cost effectiveness ratio (ICER) and cost-effectiveness acceptability curve (CEAC).

5.4 Developing a Methodology

The proposed method of identifying and measuring process utility involves adding a seventh "process" domain to the 6-domained SF-6D. Within this thesis this revised 7-domained measure will be termed the SF-7D. The method for developing the SF-7D and valuing process utility has 4 key steps:

1. Defining process utility: The identification of the most appropriate aspect(s) of the PEQ and/or CSQ to use to represent "process" and incorporate into the SF-6D;
2. The classification and description of the SF-6D plus process utility (SF-7D) health states;

3. Valuation study to empirically compare the SF-6D and SF-7D;
4. Application in practice: relating the findings back to the ePAQ RCT.

The remainder of this chapter concentrates on the initial 2 steps, and describes the design of the valuation study in step 3 in depth. Step 4 which relates to the results and application back to the ePAQ EEACT is reported in Chapter 6.

5.4.1 Defining Process Utility

The aim of this study is to improve the relevance of the ePAQ trial based outcomes by determining the measure of process utility gained or lost when completing ePAQ online and then receiving a telephone consultation, and comparing this to the online completion of ePAQ followed by a face-to-face consultation. In order to achieve this, questionnaires used within the ePAQ trial will be utilised. Within the trial patients were asked to complete the SF-12, PEQ and CSQ-8. The SF-12 is a multidimensional generic measure of health related quality of life which is widely used in clinical trials. It is a shortened version of the SF-36, and produces two summary scores, the physical and mental health scores (PCS and MCS). The SF-12 can be converted into a 6-dimensional preference-based measure of health (SF-6D), which can be used to estimate the QALY (Brazier and Roberts, 2004). This was used as the generic preference based measure, which the process domain was bolted onto.

The PEQ is a brief questionnaire which was developed in primary health care and is used for measuring patients' experience of interaction, emotion and consultation outcome (Steine et al. 2001). The questionnaire consists of 18 items on 5 dimensions: communication; emotions; short term outcome; barriers; and relations with auxiliary staff. Scoring involves sum scales for each of the 5 dimensions. Four of the scales run from 1 to 5 and the emotion scale runs from 1 to 7. A high score represents a good communication experience, positive emotions, positive consultation outcome, no communication barriers and good relations with the staff. For the ePAQ RCT only 4 dimensions were included. The respondents' relations with auxiliary staff were deemed irrelevant and were therefore omitted.

The CSQ-8 is a questionnaire which is used to measure satisfaction with services received by individuals and families (Larsen et al. 1979). Scoring of the CSQ involves an unweighted- summation of the direction-corrected response values and a calculation of measures of central tendency of the individual item ratings and for the total scale scores.

This study explores approaches to incorporating the PEQ and/or CSQ as a bolt on to the SF-6D. One approach would be to identify one question from the PEQ and CSQ captures the relevant features of process (in this case the face-to-face versus telephone consultation). This has the benefit of being directly linked to a questionnaire which has already been pre-tested for validity, reliability and

readability. However, it has the disadvantage that it is likely it will be very specific, and will also be in a different format to the SF-6D which it will be bolted onto.

A second approach is to create a new bolt-on description that links to the CSQ and/or PEQ. This may be less specific, and the link to the CSQ and PEQ may be less certain, and the phraseology less well tested. However, a format similar to the existing SF-6D could be developed. The following analysis therefore considers whether a single question or components of the PEQ/CSQ should be used to best represent the process domain.

Exploratory Analysis

Exploratory analysis was performed on the PEQ and CSQ-8 data from the ePAQ trial. Parametric, two independent sample t-tests were performed as the data were found to be normally distributed based on a comparison of mean and median domain scores (PEQ) and total scores (CSQ-8) (Table 14). The presence of a significant difference between arms as was assumed to reflect a difference induced by the process of care, as the key difference between the trial arms was the mode of consultation (face-to-face versus telephone).

Table 14 shows the results of the t-tests where there was a significant difference between trial arms. There was no reliable evidence of a statistically significant difference in the short term PEQ dimension or the CSQ total score between the control and intervention groups. There was evidence to suggest a statistically

significant difference between the control and intervention groups for the other three dimensions of the PEQ (communications, emotions and barriers).

There was also evidence to suggest a statistically significant difference between nine of the 17 PEQ items. There was no significant difference between CSQ-8 items. These findings suggest that, any of the domains, or items for which there was found to be a statistically significant difference between trial arms could potentially be used to represent the process domain. Further exploratory work was undertaken to explore this further.

Principal Component Analysis (PCA).

In order to select which component(s) of the PEQ and CSQ-8 questionnaires should be selected to represent process, initially a factor analysis was used to confirm the factor structure of the two questionnaires and the validity of the domain scores.

Factor analysis is a methodology used to understand relationships between variables. It is based on the concept that multiple observed variables have similar patterns of responses, because they are all associated with the same underlying latent variable (Fayers & Machin, 2000). It can be used to reduce the number of variables, or identify groups or clusters of variables within a dataset.

Initially a correlation matrix is computed for all variables. This shows the interrelationships between the PEQ and CSQ-8 items. Next, the number of factors

necessary to represent the data, and the method of calculating them is determined (factor extraction). PCA was used. This is a method where linear combinations of the observed variables within the existing dataset are formed, and there is no extrapolation beyond the trial sample. This compares to factor analysis where a mathematical model is derived, from which factors are estimated (Yong and Pearce, 2013).

The factors are then transformed to make them more interpretable through rotation. This simplifies the structure of the analysis to one in which each variable loads on to as few factors as possible. Orthogonal rotation was used to ensure independence between factors. The Varimax method was selected as it attempts to maximise the variance of loadings within factors, and therefore results in the smallest number of exploratory variables, providing more interpretable clusters of factors than other methods of orthogonal rotation (quartimax and equamax) (Fayers & Machin, 2000). The scores for each factor are computed.

The Kaiser-Meyer-Olkin (KMO) measure of .910 indicated that the sample size was adequate to run the analysis. All KMO values for individual variables (Appendix 13) were >.830 which was well above the acceptable limit of .5 (Field, 2009). Bartlett's test of sphericity X^2 (276) = 2241.357, p<.001, was also significant, indicating relationships between the variables included in the analysis.

As described by Field (2009), only those factors which are deemed statistically important are retained in a PCA. Therefore eigenvalues and variance values of factors were reviewed. Eigenvalues are a measure of how much variance of the data each factor explains, based on Kaiser's criterion those with an eigenvalue of greater than 1 are retained. (Field, 2009).

Table 15 shows the eigenvalues for all variables, 4 of which exceed 1, and account for 71.5% of common variance. Kaiser's criterion is said to be accurate where there are less than 30 variables and communalities after extraction are greater than 0.7 or when the sample size exceeds 250 and the average communality is greater than 0.6. In this study, there were 24 variables, and the mean communality was .715. Therefore accuracy can be assumed. Table 16 shows the factor loadings after orthogonal rotation. These indicate the importance of the variables to each factor. The significance of a factor loading was set at 0.4 as recommended by Fayers & Machin, (2000).

The CSQ-8 aims to assess client satisfaction and the results showed that all of the factors are related to each other. The PEQ is divided into three main domains: an outcome domain; a communication domain; an emotions domain. The factor loadings shown in Table 16 also supported this structure.

Table 14. PEQ domain and CSQ total scores for ePAQ trial arms.

	Control				Intervention					95% CI		
	N	Median	Mean	SD	N	Median	Mean	SD	Mean Diff	Lower	Upper	P-value
PEQ Dimension												
Communications	85	4	4.00	.75	85	4.00	4.23	.63	.37	.16	.58	.001
Emotions	85	4.5	4.69	1.27	81	5.25	5.32	1.18	.63	.26	1.01	.001
Barriers	87	4.00	4.12	.68	85	4.25	4.35	.58	.23	.04	.42	.018
PEQ Item												
Consultation Doctor	89	2	2.10	1.01	87	1	1.28	.64	-.83	-1.08	-.57	.000
Had a good talk	87	4	3.85	.84	85	4	4.28	.68	.43	.20	.66	.000
Felt reassured	87	4	3.71	.82	86	4	4.08	.79	.37	.13	.61	.003
Clinician understood what was on my mind	88	4	3.80	.89	86	4	4.27	.69	.47	.24	.71	.000
Felt was taken care of	87	4	4.02	.73	87	4	4.26	.69	.24	.03	.46	.026

	Control				Intervention					95% CI		
	N	Median	Mean	SD	N	Median	Mean	SD	Mean Diff	Lower	Upper	P-value
	Control				Intervention					95% CI		
	N	Median	Mean	SD	N	Median	Mean	SD	Mean diff	Lower	Upper	P-value
Bit difficult to connect with clinician	88	4	3.98	.96	85	4	4.25	.79	.27	.01	.53	.045
Relieved - Worried	86	5	4.80	1.49	86	6	5.50	1.24	.70	.29	1.11	.001
Cheerful - Sad	85	4	4.55	1.52	81	6	5.37	1.47	.82	.36	1.3	.001
Relaxed - Tense	85	5	4.71	1.60	84	5	5.25	1.42	.54	.09	1.00	.020

Table 15. Eigenvalues associated with each component before and after extraction, and after rotation.

Compo-nent	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.488	47.868	47.868	11.488	47.868	47.868	5.341	22.255	22.255
2	2.399	9.994	57.862	2.399	9.994	57.862	5.286	22.025	44.280
3	1.999	8.327	66.189	1.999	8.327	66.189	3.374	14.058	58.338
4	1.286	5.360	71.549	1.286	5.360	71.549	3.171	13.210	71.549
5	.859	3.580	75.129						
6	.713	2.971	78.099						
7	.575	2.397	80.496						
8	.548	2.282	82.778						
9	.482	2.007	84.786						
10	.428	1.782	86.568						
11	.415	1.730	88.298						
12	.362	1.510	89.808						
13	.355	1.479	91.287						
14	.315	1.314	92.601						
15	.293	1.223	93.823						
16	.253	1.054	94.878						

Compo-nent	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
17	.222	.926	95.803						
18	.207	.861	96.665						
19	.201	.839	97.504						
20	.162	.677	98.181						
21	.137	.571	98.752						
22	.128	.533	99.284						
23	.094	.390	99.675						
24	.078	.325	100.000						

Table 16. Rotated component matrix.

	Component			
	1	2	3	4
CSQ: Rate service received	.400	**.635**	.115	.044
CSQ: Kind of service wanted	.336	**.724**	.179	.079
CSQ: Extent service met needs	.203	**.715**	.277	.187
CSQ: Would recommend service	.164	**.820**	.159	.089
CSQ: Satisfied with amount of help	.333	**.775**	.189	.147
CSQ: Services helped to deal more effectively with problems	.044	**.617**	.283	.430
CSQ: Overall, satisfied with service	.298	**.759**	.237	.105
CSQ: If were to seek help again, would come back to services	.265	**.855**	.145	.102
PEQ: Know what to do to reduce pelvic floor problems	.084	.134	.123	**.869**
PEQ: Know what to expect from now on	.319	.074	.332	**.639**
PEQ: Able to handle pelvic floor problems differently	.054	.167	.051	**.853**
PEQ: Lead to fewer or prevent pelvic floor problems	.109	.107	.034	**.844**
PEQ: Had a good talk	**.833**	.213	.127	.121
PEQ: Felt reassured	**.746**	.236	.333	.233
PEQ: Clinician understood what was on my mind	**.778**	.245	.271	.232
PEQ: Felt was taken care of	**.804**	.260	.155	.224
PEQ: Bit difficult to connect with clinician	**.725**	.295	.244	.095
PEQ: Too much time spent on small talk	**.646**	.297	.152	-.111
PEQ: Bit difficult to ask questions	**.651**	.210	.393	.029
PEQ: Important decisions made over head	**.564**	.270	.308	.112
PEQ: Relieved – Worried	.299	.317	**.757**	.173
PEQ: Cheerful – Sad	.321	.210	**.768**	.119
PEQ: Strengthened - Worn out	.390	.284	**.766**	.122
PEQ: Relaxed – Tense	.269	.280	**.774**	.142

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

Internal Consistency Reliability tests

Following the PCA, internal consistency reliability tests were performed on the items included within the four components. This study focused on internal consistency reliability, and used 2 methods: Cronbach's α and item total correlations. Cronbach's α is a measure of internal consistency, measuring how closely related a set of items in a group are. It assumes individual items (e.g. PEQ or CSQ-8 items), or sets of items, should produce results consistent with the overall questionnaire (Fayers & Machin, 2000). Within this study, these internal consistency reliability tests were used to identify which item or domain would retain the greatest amount of underlying variance, and therefore be most the most appropriate choice for the representation of process for the process domain. Item-total correlations refer to the correlation between the items, and the total questionnaire score. For a reliable scale, all items would correlate with the total (Field, 2009). Within this study, item total correlations were used as an additional method for selecting which items/domains should represent process, where the most appropriate is those which have a high correlation with the questionnaire total. These internal consistency reliability tests were therefore undertaken on CSQ items, and PEQ items and domains (Outcome, communication and emotions). Each is discussed in turn.

Client Satisfaction (CSQ).

The internal consistency reliability statistics for the CSQ are reported in Table 17. The corrected item-total correlation refers to the correlations between each item and the total score from the questionnaire. All items in the CSQ had item-total

correlations above 0.3 which indicates a reliable scale (Field, 2009). The most highly correlated question was "CSQ: If were to seek help again, would come back to services", with a value of .864. The overall Cronbach's α for client satisfaction was .934. Table 17 shows the values of overall Cronbach's α if each item was not included in the calculation. For all of the items except for one if they were removed from the calculation Cronbach's α reduced, although all only marginally. However, if "CSQ Services helped to deal more effectively with problems" was removed Cronbach's α remained at .932. This item could therefore be removed and have no effect on the reliability of the questionnaire. Within this study we assume that those items or domains which have the biggest effect on the reliability of the questionnaire when removed would be the most appropriate choice to represent process. Within the client satisfaction domain removing "CSQ: If were to seek help again, would come back to services" had the greatest effect, reducing Cronbach's α to .916.

Table 17. Item-total statistics for client satisfaction.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
CSQ: Rate service received	23.55	13.648	.682	.513	.929
CSQ: Kind of service wanted	23.63	13.516	.747	.593	.924
CSQ: Extent service met needs	23.85	12.795	.769	.630	.923
CSQ: Would recommend service	23.50	13.428	.787	.674	.921
CSQ: Satisfied with amount of help	23.56	12.809	.814	.721	.919
CSQ: Services helped	23.90	13.317	.649	.489	.932

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
to deal more effectively with problems					
CSQ: Overall, satisfied with service	23.62	12.918	.829	.708	.918
CSQ: If were to seek help again, would come back to services	23.47	13.153	.864	.793	.916

Outcome (PEQ)

For the outcome domain, item-total correlation values were again all over .3 and are therefore reliable, although values for all domain items were lower than for the client satisfaction domain. The item showing the greatest correlation was "PEQ: know what do to reduce pelvic floor problems." This had an item-total correlation of .783.

The overall Cronbach's α for PEQ outcome was .864. Table 18 shows the values of overall Cronbach's α if each item was not included in the calculation. Deleting the same PEQ response "Know what to do to reduce pelvic floor problems" would have the biggest impact on the overall Cronbach's α, reducing it to .797.

Table 18. Item-total statistics for outcome.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
PEQ: Know what to do to reduce pelvic floor problems	**8.79**	**11.131**	**.783**	**.623**	**.797**
PEQ: Know what to expect from now on	**8.12**	**12.950**	**.594**	**.362**	**.873**
PEQ: Able to handle pelvic floor problems differently	**9.02**	**11.423**	**.760**	**.606**	**.807**
PEQ: Lead to fewer or prevent pelvic floor problems	**8.86**	**11.760**	**.719**	**.538**	**.824**

Communication (PEQ).

Item-total correlation values again all exceeded .3 and were therefore deemed to be reliable. The item showing the greatest correlation was "PEQ: Clinician understood what was on my mind" This had an item-total correlation of .822.

The overall Cronbach's α for PEQ communication was .930. Table 19 shows that the Cronbach's α values associated with the removal of any communication item were very similar and remained high for all items, indicating all items within this domain were important. Removal of "PEQ: clinician understood what was on my mind" had the greatest impact on the reliability, reducing Cronbach's α to .916. This

was also a question which showed a significant difference in responses between the ePAQ trial arms.

Table 19. Item-total statistics for communication.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
PEQ: Had a good talk	29.11	19.446	.787	.719	.918
PEQ: Felt reassured	29.27	19.174	.799	.712	.917
PEQ: Clinician understood what was on my mind	29.13	19.103	.822	.717	.916
PEQ: Felt was taken care of	29.02	19.873	.812	.723	.917
PEQ: Bit difficult to connect with clinician	29.07	18.857	.796	.657	.918
PEQ: Too much time spent on small talk	28.84	21.518	.655	.466	.928
PEQ: Bit difficult to ask questions	29.08	19.024	.745	.578	.922
PEQ: Important decisions made over head	28.83	20.940	.674	.546	.927

Emotion (PEQ)

Corrected item-total correlation values were all above .3, indicating that all items correlated well with the overall score. The item with the greatest value was the PEQ: Strengthened – worn out, with a value of .787.

The overall Cronbach's α was .895 for the emotion domain. Table 20 shows the values of overall Cronbach's α if each item is excluded in the calculation. For all items if they were removed Cronbach's α reduced, but again only marginally.

Removing "PEQ: strengthened-worn out" had the greatest impact on the reliability, reducing Cronbach's α to .861.

Table 20. Item-total statistics for emotion.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
PEQ: Relieved – Worried	14.87	15.049	.777	.611	.862
PEQ: Cheerful – Sad	15.05	14.422	.747	.575	.874
PEQ: Strengthened - Worn out	15.07	15.759	.787	.641	.861
PEQ: Relaxed – Tense	15.02	14.381	.773	.631	.863

A summary of the key findings from both the exploratory statistics and the PCA are reported in Table 21.

Table 21. Summary of exploratory analysis and PCA.

Questionnaire	Item	Findings
CSQ	Services helped to deal more effectively with problems	Cronbach's α: if removed had no effect on reliability of questionnaire
	If were to seek help again, would come back to services	Cronbach's α: if removed had largest effect on reliability of questionnaire. Highest item-total correlation within domain (.864)
PEQ	Know what to do to reduce pelvic floor problems	Cronbach's α: if removed had largest effect on reliability of questionnaire. Also highest item-total correlation within domain (.783)
	Know what to expect from now on	Cronbach's α: if removed had no effect on reliability of questionnaire

Questionnaire	Item	Findings
	Had a good talk	Identified in exploratory analysis, as showed statistically significant difference between trial arms. Is also directly related to consultation type.
	Clinician understood what was on my mind	Identified in exploratory analysis, as showed statistically significant difference between trial arms. Cronbach's α: if removed had largest effect on reliability of questionnaire. Also highest item-total correlation within domain (.822)
	Cheerful – Sad	Identified in exploratory analysis, as showed statistically significant difference between trial arms.
	Strengthened - Worn out	Cronbach's α: if removed had largest effect on reliability of questionnaire. Also highest item-total correlation within domain (.787)

The findings from the exploratory analysis, and PCA identified two potential approaches to representing the process domain:

1. If the "PEQ: The clinician understood what was on my mind" question was removed, it would have the largest effect on the reliability of the questionnaire. This question also displayed the highest value of item-total correlation, and was identified in the exploratory analysis as one of the 5 items with a significant difference between treatment arms.
2. The communication domain of the PEQ was also identified as important. The overall Cronbach's α was .930. Cronbach's α values associated with the

removal of any communication item were very similar and remained high for all items, indicating all items within this domain were important.

Due to these findings, and because a priori it was thought the main benefit for the technology would be in the communication between the patient and the doctor, these 2 potential approaches were explored further in a pilot study.

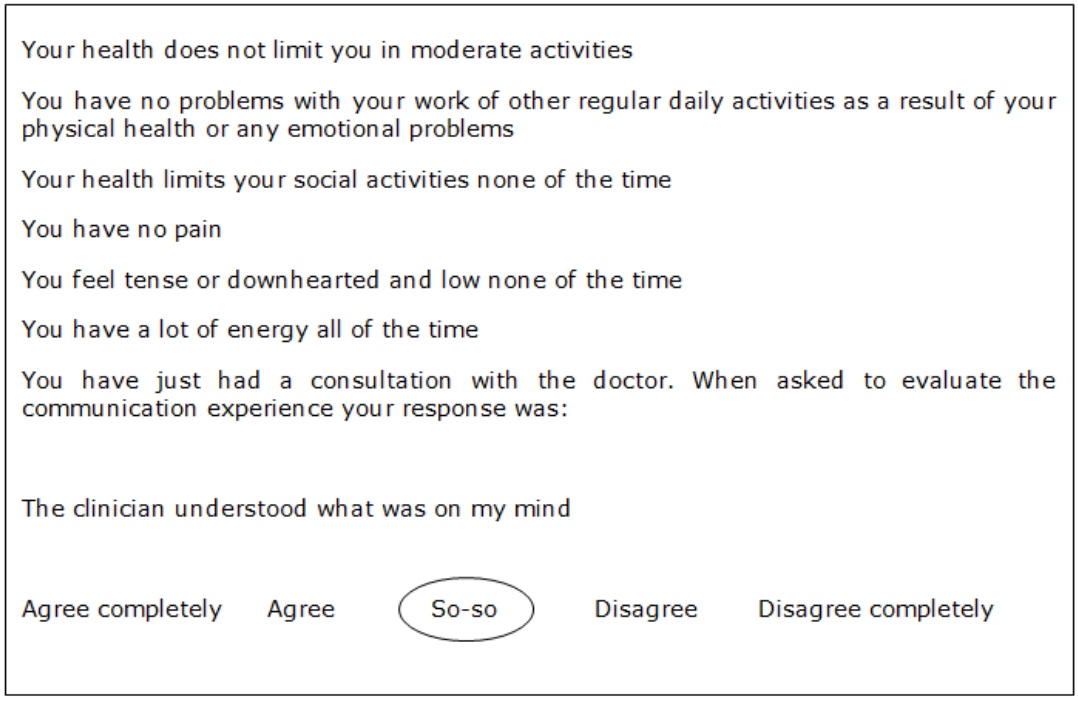
5.4.2 The classification and description of the SF-6D plus process utility (SF-7D) health states.

Pilot study

Method

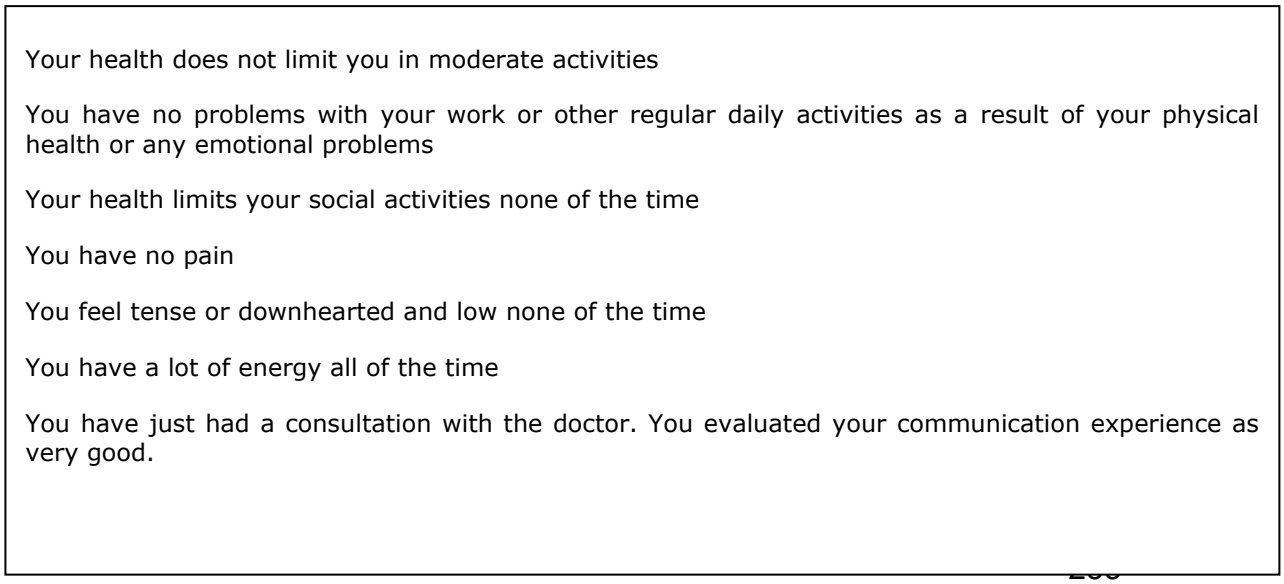
The two alternative approaches to adding the process domain to the SF-6D which were highlighted previously were compared. In one approach, the single PEQ question was added as the bolt-on. As can be seen in Figure 16 this produces an SF-7D health state with a process domain which looks very different to the SF-6D domains.

Figure 16. S F-7D Health state using methodology A



In the second approach, communication as a whole was described on a bespoke scale that was more closely matched to the SF-6D (Figure 17). The scale included 5 levels: very poor; poor; average; good, and very good.

Figure 17. SF-7D Health State using methodology B.



If the latter approach were used (Figure 17) an additional step would be required, as the new scale (very poor, poor, average, good, very good) would need to be linked to the PEQ communication questions. Therefore, the pilot study included:

- Health states with a single question to represent the process domain (Figure 16)
- Health states with a bespoke process description (Figure 17)
- A set of PEQ communication health states, which respondents were asked to rate using the new scale (Figure 18).

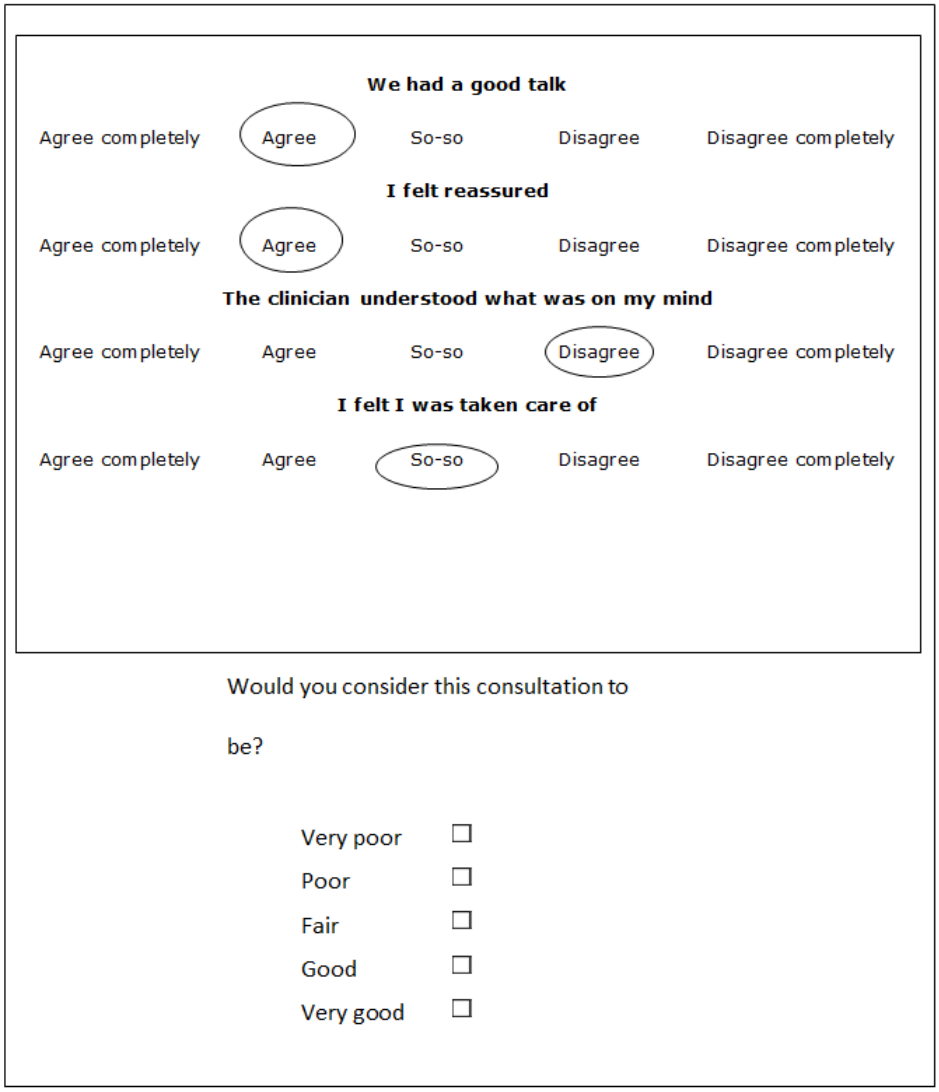
The aim of the pilot study therefore was:

1 To determine which approach respondents preferred for the representation of the process domain: a single question, or a bespoke scale.

2 To determine whether it was possible to ask respondents to rate the PEQ health state, and calculate a rating score for each combination of PEQ responses.

The pilot study was therefore performed to test the formatting of the health and process state, and check it was feasible to ask respondents to rate PEQ health states. The findings would be used in the design of the valuation study. Therefore, it was decided that, although for the valuation study itself a sample representative of the population would be preferred – as this would provide utility estimates, for the pilot study a convenience sample was adequate to inform the study design.

Figure 18. **Rating a PEQ communication health state**



All individuals working for RTI-Health Solutions and based in the UK offices were sent questionnaire booklets (Appendix 14) and asked if they would like to participate in the pilot study.

The interview booklets contained the 3 components:

1 Respondents were shown 3 SF-7D health states (e.g. Figure 16) : Methodology A). The process domain was represented by one question taken directly from the PEQ questionnaire. They were then asked to value these health states using the VAS.

2 Respondents were shown 3 further SF-7D health states (e.g. Figure 17: Methodology B). This time the process domain was represented as a rating score of: very good; good; average; poor; very poor. Respondents were again asked to value these health states using the VAS.

Respondents were then asked which set of descriptions they preferred: Methodology A or Methodology B.

3 Respondents were shown 4 scenarios, which were based on the communication domain of the PEQ (e.g. Figure 18). They were informed that the scenarios described their feelings immediately after a consultation with the doctor. For each scenario they were asked if they would consider the consultation to be very poor, poor, fair, good or very good.

Results.

Twenty-three individuals returned completed questionnaire booklets. All respondents had valued all health states using both approaches, indicating that

either approach was a potential. All respondents were also able to rate the PEQ scenarios using the bespoke scale.

Of the 23 respondents, 8 preferred Methodology A where the question was taken directly from the PEQ questionnaire, and 15 preferred methodology B where the bespoke rating scale was used.

Conclusion

The preferred methodology for the representation of the SF-7D health state was methodology B. Therefore, based on the Mann-Whitney test results, and the findings from the pilot study, the communication domain as a whole was selected to represent the consultation experience. This introduced the additional requirement: that the responses to the PEQ communication domain can be converted to an overall rating for the domain.

5.4.3 Selection of the SF-6D and SF-7D health states

Patients included in the ePAQ EEACT were asked to complete the SF-12 at baseline (sub-set of patients) and 6 months follow-up. The ultimate aim of the process utility valuation study was to relate the findings back to the ePAQ EEACT. Therefore the initial selection criterion for health states was that they must be those experienced by ePAQ trial patients. There were 252 health states experienced by patients within the trial. A review of the health states indicated that majority of patients' experienced less severe health states.

The second selection criterion mirrors Krabbe et al. (1999) who selected states purely to ensure an even mix of good, moderate and bad health states based on the EQ-5D value. The health states experienced ranged from 111112 (utility estimate of 0.922) to 34555 (utility estimate of 0.345). Therefore, health states included were also within this range. The most commonly reported health state was 111112 (n=11). A total of 20 health states were selected.

The seventh attribute was added to all of these 20 health states, again selection was based on methodology used by Krabbe et al. (1999) whereby levels were chosen randomly, occasionally avoiding a level too unlikely in combination with the other six levels. A further two SF-7D health states were added differing by only one level in the process domain from an already selected SF-7D state. The resulting 20 SF-6D and 22 SF-7D health states can also be seen in Table 22. An additional "dead" health state was also included. Therefore, the final number of included health states was 21 SF-6D and 23 SF-7D.

Table 22. Selected SF-6D Health States

Health State	SF-6D	Observed Frequency in trial	SF-7D
111112	0.922	11	111112(1)
-	-	-	111112(2)
111113	0.922	5	111113(1)
111212	0.922	4	111212(1)
111122	0.863	6	111122(3)

Health State	SF-6D	Observed Frequency in trial	SF-7D
221112	0.859	1	221112(3)
121122	0.8	3	121122(3)
-	-	-	121122(4)
221122	0.8	3	221122(2)
241212	0.782	1	241212(5)
241224	0.723	2	241224(1)
131132	0.722	4	131132(4)
343112	0.671	1	343112(3)
132133	0.66	4	132133(4)
143233	0.657	3	143233(5)
243223	0.657	1	243223(2)
243334	0.615	4	243334(2)
243434	0.58	1	243434(5)
343423	0.535	1	343423(2)
343435	0.507	2	343435(5)
345535	0.42	1	345535(4)
345555	0.345	2	345555(5)

5.4.4 Online valuation study to empirically compare the SF-6D and SF-7D

Developing the survey.

The online survey was developed with the assistance of Epigenesys, a university spin-off software development company based in Sheffield, UK, and was funded through the ScHARR data collection fund. The system used for the delivery of health state valuation tasks had already been developed by staff at ScHARR in conjunction with Epigenesys. This was updated to include VAS questions. An accredited online participant panel SSI was used to source the respondents.

The survey initially included project information and consent to take part. It then consisted of three parts:

1. **The SF-6D/SF-7D health states:** Respondents were shown either a set of 21 SF-6D health states, or a set of 23 SF-7D health states, and were asked to value them using the VAS. Despite SG or TTO approaches being the preferred approach for the valuation of health states, the decision was made to use the VAS within this study. Although the VAS is not preferred by NICE, for reasons discussed in chapter 2 (e.g. it is not a choice based method), it does provide the simplest approach to valuing health states, and is still a recognised approach (Green et al, 2000). It was therefore selected for 2 main reasons: firstly, to reduce the burden of completion, the questionnaire included 3 sections in total, the first with either 21 or 23 questions, the second with 25 and the third with 7 questions. Asking respondents to value the health

states using either TTO or SG would therefore have added too much complexity and time for completion to the study. And secondly, because this is a feasibility study for process utility. It is accepted that taking this approach means that the SF-6D, which is based on valuations using SG, will be valued differently to the SF-7D, and that different valuation methods will produce different results, with SG scores generally being greater than VAS scores (Bass et al. 1994; Read et al. 1984). Any further work should use the same valuation method for the main instrument and bolt-on, or alternatively, the valuations of the full instrument plus the bolt-on should be undertaken at the same time.

2. **The PEQ communication scenarios:** The findings from the exploratory analysis and pilot study indicated that the communication domain as a whole should be used to represent the consultation experience, using the format where the domain and its severity is reflected by the sentences: "You have just had a consultation with your doctor. You valued your communication experience as very good." The ratings ranged from very poor to very good (very poor; poor; average; good; very good). Taking this approach, the ePAQ EEACT patients' responses to the PEQ communication domain needed to be converted into the rating scale responses using multinomial regression analysis. In order to do this there was a requirement to obtain PEQ communication domain responses and associated rating scores from a larger sample of respondents. The co-efficients from the regression analysis could then be applied to the ePAQ PEQ communication responses to obtain ratings for each trial participant. The PEQ communication domain has 4 items, each

with 5 potential responses. This equates to 625 potential PEQ communication states. It is impractical to collect enough data on each to determine the relationship between the PEQ responses and overall rating for the health state. Therefore, an orthogonal fractional factorial design was used to select the number and combination of health states which represent the necessary interactions between responses to provide meaningful outcomes. This resulted in a set of 25 PEQ communication states. Respondents were asked whether they would consider each communication experience to be very good, good, average, poor or very poor.

3. **Demographic questions:** Respondents were asked questions, including: age; gender; ethnicity; marital status; education.

Sample

The study aimed to administer the survey to 200 members of the UK general population. This sample size was based on Krabbe et al. (1999) and budget constraints. Krabbe et al. (1999) sent out 185 paper questionnaires and achieved a 47% response rate, resulting in 87 valid questionnaires being returned. They discuss their sample size as a key limitation to their study. This study therefore aimed to address that limitation.

Sample Recruitment

Participants were randomly selected from those in the panel eligible to take part in this study, and who fulfilled the age and gender quotas that were equivalent to the UK general population. All members of the panel had completed a double opt-in

process to join the panel and have consented to completing online surveys over the course of their membership of the panel. The participants randomly selected from the internet panel were emailed with a link to take part in the survey.

Data Analysis

All analyses were performed using SPSS. Analysis included demographic characteristics of the two groups of participants. Means and standard deviations for all valued SF-6D and SF-7D health state descriptions. Statistical testing within pairs (SF-6D versus SF-7D) was performed using t tests.

Descriptive analysis was used to explore the relationship between the level of the added domain and the decrement. Further exploratory analysis was then performed to explore the relationship between the decrement and the severity of the health state.

Regression analysis was run on the PEQ scenarios and rating scores. The co-efficients were applied to the ePAQ trial patients to obtain a rating for each participant's combination of PEQ communication responses. An ICER and CEAC were then reported including process.

5.5 Conclusion

This chapter reported the development of the methodology for a health state valuation study which added an additional process domain on to the SF-6D. It began with a literature review of bolt-on studies, which provided an overview of the current status of bolt-on studies and their methods. The study design was informed

by these methods, but also introduces new approaches, making this valuation study novel for several reasons. The process domain is represented using components of an existing and established instrument which has undergone psychometric testing for reliability and validity. The approach negates the need to develop hypothetical health states, and their associated limitations (E.g. their link with clinical practice, assurance of the validity and reliability of health states). It also allows the study findings to be related back to the ePAQ trial. This has the advantage of reducing the focusing effects reported by both Krabbe et al. (1999) and Brazier et al. (2011), and also increasing internal validity.

The online completion of the valuation study differs from Krabbe et al. (1999) who sent out postal questionnaires. This involved working with a software development company to develop online VAS questions and allowed the survey to be completed by a larger sized sample, more comparable to the general population.

This chapter has provided a detailed description of key components of the process utility valuation study design. Chapter 6 provides further detail on the methodology, and then reports the study results in detail.

5.6 References

- Alava, MH. Brazier, J. Rowen, D. Tsuchiya, A (2013) Common Scale valuations across Different Preference-Based Measures: Estimation Using Rank Data. *Medical Decision Making,* 33(6)839-852

- Bass, EB. Seinberg, EP. Pitt, HA. Lillemoe, KD. Saba, GP. Johns, C. (1994) Comparison of the rating scale and the standard gamble in measuring patient preferences for outcomes of gallstone disease. *Med Decision Making,* 14(4),307-14.

- Brazier, J. Rowen, D. Tsuchiya, A. Yang, Y. Young, T. (2011) The impact of adding an extra dimension to a preference-based measure. *Social Science and Medicine*, 73,245-253.

- Brazier, JE. Roberts, JR. (2004) The estimation of a preference-based index from the SF-12. *Medical Care*, 42, 851-859.

- Fayers, PM and Machin, D (2000) Quality of Life. Assessment, analysis and interpretation. Wiley, England.

- Field, A. (2009) Discovering Statistics using SPSS. Third Edition. Sage Publishing, USA.

- Green, C. Brazier, J. Deverill, M. (2000) Valuing Health-related quality of life. A review of Health State Valuation Techniques. *Pharmacoeconomics*, 17(2), 151-165.

- Krabbe, P. Stouthard, M. Essink-Bot, M. Bonsel, G. (1999) The effect of adding a cognitive dimension to the EuroQol Multiattribute Health-Status Classification System. *J Clin Epidemiol,* 52(4), 293-301.

- Larsen, DL. Attkisson, CC. Hargreaves, WA. and Nguyen, TD. (1979) Assessment of client/patient satisfaction: Development of a general scale, *Evaluation and Program Planning*, 2, 197-207

- NICE (2013) National Institute for Health and Care Excellence. Process and Methods Guides: Guides to the methods of technology appraisal 2013. Available from: http://publications.nice.org.uk/pmg9 [Accessed 28th August 2015].

- Read, J. Quinn, R. Berwick, D. Fineberg, H. Weinstein, M. (1984). Preferences for health outcomes – comparisons of assessment methods. Medical Decision Making, 4, 315-29.

- SiegeI, JE. Torrance, GW. Russell, Luce, BR. Weinstein MC. Gold, MR. Members of the Panel on Cost Effectiveness in Health and Medicine. (1997) Guidelines for Pharmacoeconomic Studies. Recommendations from the Panel on Cost Effectiveness in Health and Medicine. *Pharmacoeconomics,* 11(2), 159-168.

- Steine, S. Finset, A. Laerum, E. (2001) A new, brief questionnaire (PEQ) developed in primary health care for measuring patients' experience of interaction, emotion and consultation outcome. *Family Practice* 18,410-418.

- Swinburn, P. Lloyd, A. Boye, KS. Edson-Heredia, E. Bowman, L. Janssen, B. (2013) Development of a disease-specific version of the EQ-5D-5L for use in patients suffering from psoriasis: lessons learned from a feasibility study in the UK. *Value in Health, 16,1156-1162.*

- Wolfs, C. Dirksen, C. Kessels, A. Willems, D. Verhey, F. Severens, J. (2007) Performance of the EQ-5D and the ED-5D+C in elderly patients with cognitive impairments. *Health and Quality of Life Outcomes,* 5,33.

- Yang, Y. Brazier, J. Tsuchiya, A. (2008) The effect of adding a 'sleep' dimension to EQ-5D. Health Economists' Study Group meeting January.

- Yang, Y. Rowen, D. Brazier, J. Tsuchiya, A. Young, T. Longworth, L. (2015) An exploratory study to test the impact on three “bolt-on” items to the EQ-5D. *Value Health* 18(1), 52-60.

- Yong, AG. and Pearce, S. (2013) A Beginner’s Guide to Factor Analysis. *Tutorials in Quantitative Methods for Psychology*. 9(2),74-94.

Chapter 6: Process Utility Bolt-on Study Results and Discussion.

This chapter describes the process utility bolt-on study analysis in depth. It begins by providing an overview of the study sample, and then outlines both the descriptive and exploratory analysis undertaken. The findings are then applied to the ePAQ trial, and subsequently an updated ICER and CEAC is reported.

6.1 Results

6.1.1 Demographic Data

Of the 609 respondents who began the survey, 268 completed, resulting in a dropout rate of 66%. Of the 268 who completed, 8 were excluded for the following reasons: 3 provided responses deemed to be irrational; 1 valued all health states as 100; 1 valued all health states as 0; 1 valued dead as 100; 2 valued dead as 0 and all other health states as 100.

Of the remaining 260, 130 completed valuations of the 21 SF-6D health states, and 130 completed valuations of SF-7D health states. The background characteristics of these 260 respondents can be seen in Table 23.

Table 23. Background characteristics.

Background characteristics	National Census data (2011)[9]	SF-6D (n=130)	SF-7D (n=130)
Age (years)	%	n(%)	n(%)
18-39	40.5	34(26.1)	38(29.2)
40-59	32.4	33(25.4)	32(24.6)
60-75+	27.1	63(48.5)	60(46.2)
Gender	%	n(%)	n(%)
Male	49	71(54.6)	63(48.5)
Female	51	59(45.4)	67(51.5)
Education level[10]		n(%)	n(%)
O levels/GCSCs and equivalent	-	56 (43.1)	60(46.2)
Degree and higher degree	-	43(33.1)	37(28.5)
Professional qualifications/other vocational related qualifications	-	22(16.9)	23(17.7)
Foreign qualifications	-	2(1.5)	0(0)
No qualifications	-	7(5.4)	10(7.7)
Employment[11]status	%	n(%)	n(%)
Student	8.85	6(4.6)	4(3.1)
Employed	50.11	37(28.5)	49(35.4)
Home-maker	4.17	9(6.9)	7(5.4)
Self-employed/freelance	9.34	9(6.9)	13(10.0)
Unemployed	6.81	2(1.5)	4(3.1)
Long-term sick/disabled	3.88	8(6.2)	5(3.8)
Retired	13.10	59(45.4)	51(39.2)
Marital status	%	n(%)	n(%)
Never married & never registered in a same-sex civil partnership	34.6	39(30.0)	41(31.5)
Married/In a registered same-sex civil partnership	46.8	66 (50.8)	66(50.8)
Separated, but still legally married/separated	2.6	3(2.3)	2(1.5)

[9] Calculated for age 18-24 based on 2011 Census data. Based on 2011 Census data tables: Usual resident population by 5-year age group and sex

[10] Categories for qualification levels have altered since 2011 census. Therefore, are not directly comparable.

[11] ONS census data includes "economically inactive: Other" category, therefore % do not add up to 100.

Background characteristics	National Census data (2011)[9]	SF-6D (n=130)	SF-7D (n=130)
but still legally in a same-sex civil partnership			
Divorced/formerly in a same-sex civil partnership which is now legally dissolved	9.0	15(11.5)	13(10.0)
Widowed/surviving partner from a same-sex civil partnership	7.0	7(5.4)	8(6.2)

6.1.2 Health State Valuations

The data were then transformed, using the respondents' valuation of the "dead" health state, onto a scale where 0 represents "dead" and 1 represents "full health", using the methodology reported by Torrance, (1986), where if dead is valued at 0, the preference value for all other health states is the reported value, and where dead is not valued as the worst state, but an intermediate state *d*, then the health state preference value is calculated using the following formula:

Health state preference value = (x-d)/(1-d)

where *x* is the scale placement of the health state.

Table 24 shows the mean valuations and standard deviations for the adjusted health states, the utility decrements and the results of the Mann-Whitney test for differences between the SF-6D and SF-7D valuations. This test was chosen as data were non-normally distributed.

Significant differences ($p<.05$) between mean SF-6D and SF-7D valuations were identified in 10 comparisons (highlighted in red). A comparison of the mean valuations between the SF-6D and SF-7D health states can be seen in Figure 19.

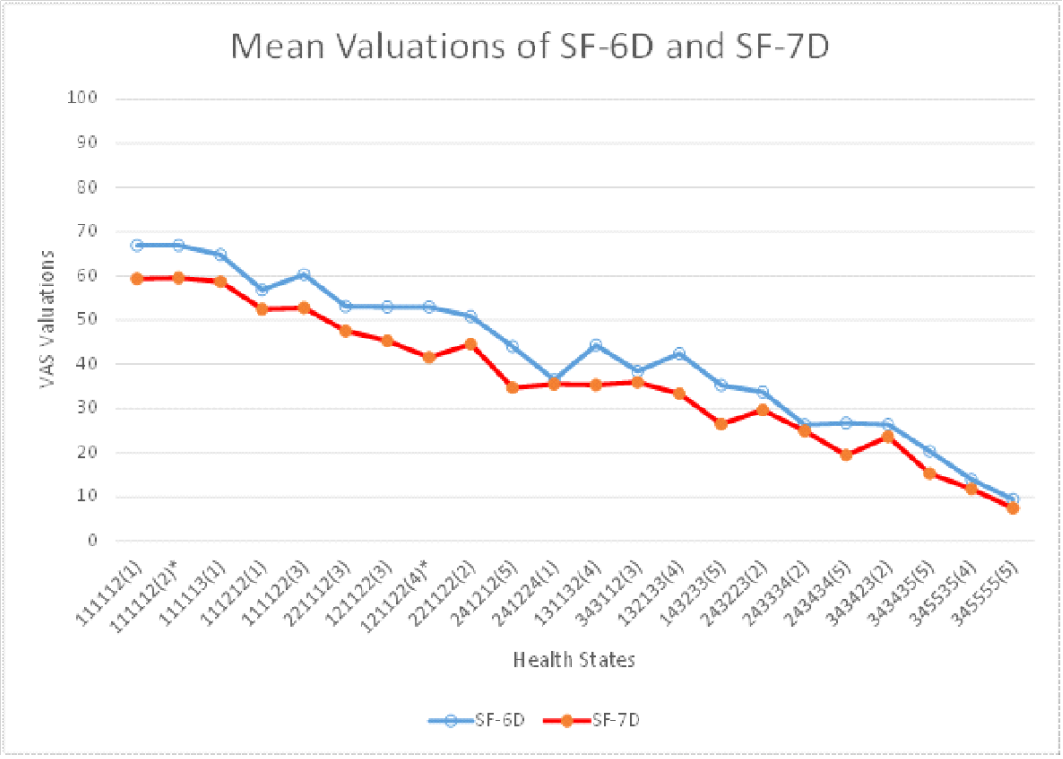
Table 24. Mean and standard deviations for SF-6D and SF-7D Decrements.

SF-6D(7D) Health states	SF-6D			SF-7D			Difference	Mann-Whitney Test
	N	Mean	SD	N	Mean	SD		p
111112(1)	130	66.88	32.20	130	59.40	33.21	-7.48	.017
111112(2) (level effects[12])	-	-	-	130	59.58	33.14	-7.30	.017
111113 (1)	130	64.82	31.66	130	58.71	34.70	-6.11	.139
111212(1)	130	56.84	31.15	130	52.45	33.19	-4.39	.217
111122(3)	130	60.27	30.26	130	52.79	30.88	-7.47	.029
221112(3)	130	53.08	27.86	130	47.54	32.44	-5.54	.263
121122(3)	130	52.97	29.06	130	45.33	31.95	-7.63	.040
121122(4) (level effects)	-	-	-	130	41.55	34.95	-11.42	.008
221122(2)	130	50.86	28.02	130	44.55	32.54	-6.30	.137
241212(5)	130	44.00	29.42	130	34.72	32.83	-9.28	.018

[12] Level effects refer to the effect of changing only the level of the process domain. Two pairs of health states were included which differed only in the level of the process domain: 111112(1) and 111112(2), 121122(3) and 121122(4).

SF-6D(7D) Health states	SF-6D			SF-7D			Difference	Mann-Whitney Test
	N	Mean	SD	N	Mean	SD		p
241224(1)	130	36.50	29.31	130	35.55	31.73	-0.94	.902
131132(4)	130	44.31	29.39	130	35.32	32.71	-8.99	.023
343112(3)	130	38.44	29.56	130	35.96	28.88	-2.48	.430
132133(4)	130	42.45	29.60	130	33.39	30.27	-9.06	.008
143233(5)	130	35.29	29.69	130	26.48	31.69	-8.80	.025
243223(2)	130	33.74	29.96	130	29.68	29.86	-4.07	.331
243334(2)	130	26.33	29.93	130	24.91	29.35	-1.42	.766
243434(5)	130	26.69	29.09	130	19.46	30.30	-7.23	.037
343423(2)	130	26.40	28.66	130	23.62	30.35	-2.78	.462
343435(5)	130	20.35	30.94	130	15.33	29.76	-5.02	.164
345535(4)	130	13.94	31.45	130	11.82	31.84	-2.12	.589
345555(5)	130	9.47	28.42	130	7.50	30.90	-1.97	.647

Figure 19. Mean Valuations of SF-6D and SF-7D (N=260)



Level effects

Level effects refer to the effect of changing only the level of the process domain. Two pairs of health states were included which differed only in the level of the process domain. These are also shown in Table 24. The addition of a “very good” process domain to SF-6D health state 111112 resulted in a decrement of -7.48. The addition of a “good” process domain reduced this decrement to -7.30. When considering the addition of more severe process domains, the addition of an “average” process domain to SD-6D health state 121122, resulted in a decrement of -7.63. This decrement increased to -11.42 when the added process domain was poor.

6.1.3 Exploring the decrements

Descriptive analysis

Descriptive analysis was used to explore the relationship between the level of the added domain and the decrement. Table 25 reports the mean decrement for each process domain level. The results indicate that as the rating declined (From 1 to 5) the value of the decrement did not reduce monotonically. The results suggest there may be a trend of increasingly greater decrements for worse process domains. Exploratory analysis was therefore performed to explore this further.

Table 25. Mean Decrements for the 5 process domain levels

Domain level	Domain De-scription	N	Minimum	Maximum	Mean	Std Deviation
1	Very good	4	-7.48	-.94	-4.73	2.82
2	Good	5	-7.30	-1.42	-4.37	2.43
3	Fair	4	-7.63	-2.48	-5.78	2.40
4	Poor	4	-11.42	-2.12	-7.90	4.01
5	Very poor	5	-9.28	-1.97	-6.46	3.01

Exploratory Analysis

In order to further explore the relationship between the decrement and the severity of the health state, the number of domain levels was reduced from 5 to 3: Poor, fair and good (by combining the two upper domains, and two lower domains, and maintaining the middle domain). The results, with this reduced number of domains, indicate that as

the rating level declined (from 1 to 3) the value of the decrement reduced monotonically (Table 26).

Table 26. Mean Decrements for the 3 process domain levels

Domain level	Domain De-scription	N	Minimum	Maximum	Mean	Std Deviation
1	Good	9	-7.48	-.94	-4.5331	2.44675
2	Fair	4	-7.63	-2.48	-5.7829	2.39775
3	Poor	9	-11.42	-1.97	-7.0989	3.33829

The decision was therefore made to continue with the analysis using the findings from the 3 domains, rather than the 5 domains. Although this approach has limitations, for example a loss of information (which may have been gained from including 5 levels), and particularly given the small sample size for each domain (between 4 and 5), it was felt that it would result in "cleaner" findings and aid in ascertaining the feasibility of undertaking further and more detailed analysis where a 5 levelled process domain could be explored with higher sample sizes for each domain level.

The descriptive analyses for the 3 process levels (Table 26) therefore show that overall there is a reduction in utility valuations associated with the addition of a process domain. Further statistical analyses were performed to investigate whether the severity of the starting health state influenced the level of the decrement.

Multiple linear regression was used to assess the association of the process domain and mean utility across the study sample. Regression was run on mean utility rather

than individual patient data, as tariff values for the SF-6D are based on mean health state regressions as reported by Brazier and Roberts (2004), and which in turn was because the use of population means are recommended for cost-effectiveness analyses.

Three separate specifications were assessed: the first estimated mean utility as a function of a constant plus process level, the second also included a measure of overall health state severity, and the third included health state severity, plus an interaction term between health state severity and process level. The model for the latter, and from which the other two models are simplifications, is defined as:

$$\text{Mean utility} = \alpha + \text{process level} + \text{severity} + \text{process level} * \text{severity} + \varepsilon$$

Where the process level is defined as a categorical variable with 3 levels (poor, fair, good) and severity is defined by characteristics of the underlying SF-6D health state. Different approaches to defining severity were considered. These included:

- The use of the tariff utility value for the underlying SF-6D health state;
- A simple count of the health state descriptors, where 111112 would be equal to a severity of 7;
- In order to simplify the analysis, in the presence of only 22 data points, severity was represented with dummy variables, where 1 = healthy patients, and 0 = sick patients. These are binary data categories determined as values falling above and below the median SF-6D utility scores obtained from the survey.

A summary of the regression statistics can be seen in Table 27. The initial model, containing only the process domain and decrement, did not reach statistical significance (p>.001). Adding severity to the model increased the predictive power of the model, as the adjusted r^2 increased from .119 to .555, and resulted in a statistically significant F-ratio (sig .000). This second model therefore significantly improved our ability to predict the dependent variable. These initial two models considered only the main effects of each of the independent variables, and so assumed additive effects. However, as previous research has indicated, (Brazier et al. 2010) interaction between these variables may exist. An interaction between severity and domain was therefore included, but this reduced the predictive power of the model slightly (adjusted r^2 = .531), however, the F-ratio remained significant.

Table 27. Summary statistics: Regression (base case)

Model	Process domain	Process domain + severity	Process domain + severity + (severity*domain)
Adjusted R2	.119	.555	.531
R2	.161	.598	.598
Sig of F-ratio	.064	.000	.001
Constant	-3.244	-.427	-.402
Process domain β	-1.283	-1.716*	-1.728
Severity β	-	-3.900**	-3.952
Interaction	-	-	.026

*Significant (p<.005); ** Significant (p<.001); β = unstandardised co-efficient.

6.1.4 Sensitivity Analysis

The analysis was re-run on one additional population including only those who provided rational responses to the level effects question - where the valuation for 111112(1) < 111112(2) and/or 121122(3) < 121122(4). This resulted in N = 186. The proportion of respondents in older age groups remained high. Of these 186 respondents, 130 completed the SF-6D and 56 completed the SF-7D. No participants were excluded from the SF-6D completions as there were no health states included to test for level effects for the SF-6D. Comparisons of mean valuations between SF-6D and SF-7D health states, using the Mann-Whitney test showed only three reaching statistical significance: 111112(2), 121122(4) and 132133(4). Mean valuations for each of the 5 domains showed that there was a monotonic relationship between the decrement and valuation. The health states were again reduced to 3, and the monotonic relationship remained.

The regression models were re-run (Table 28). This resulted in a higher adjusted r^2 for the model containing only the process domain and decrement, and lower adjusted r^2 values for the remaining 2 models, however, none of the models reached statistical significance ($p < .001$). The results did improve on the primary analysis, and therefore either respondent sample may be used.

Table 28. Summary statistics: Regression (sensitivity analysis)

Model	Process domain	Process domain + severity	Process domain + severity + (severity*domain)
Adjusted R^2	.231	.192	.150
R^2	.268	.519	.272
Sig of F-ratio	.014	.051	.119
Constant	-.864	-1.795	1.731
Process domain β	-.518	-6.444	-8.363
Severity β	-	.883	-1.508
Interaction	-	-	1.225

β = unstandardised co-efficient.

On the basis of these findings, the results from the regression analysis on the total population (base case), including the process domain, severity and the interaction term will be used in the trial analysis. This was selected as both this regression model, and the models including the process domain and severity were significant, but this model also accounted for interactions between variables which have been suggested by previous research (Brazier et al. 2010).

6.1.5 Application to the ePAQ Trial.

The findings from the process utility valuation study were applied to the ePAQ trial. The analysis was based on the NHS perspective. This allowed the willingness to pay threshold to be applied and compared to the CEAC reported in Chapter 3.

In order to apply the process utility findings to the trial data, it was necessary to determine which combinations of the PEQ communication responses from the trial corresponded with which rating score. As explained in detail in Chapter 5, 25 PEQ questions were included in the online survey. Respondents were asked to rate them as "very good, good, average, poor or very poor". Each health state was rated between 269 and 272 times. Multinomial regression was run to determine what combination of PEQ scores would relate to which rating score. The categorical predictors (factors) were the PEQ responses. There were 4 predictors for each rating scale (4 questions within the PEQ communication domain). The dependent variable was the rating score. There were no covariates, and a main effects model was run so as to include both main effects and factor-by-factor interactions.

The chi-square statistic for the main effects model was significant (<0.001), indicating that the factor by factor interactions between the PEQ responses had a significant effect on predicting the rating of the health state. Likelihood ratio tests were significant for all predictors (PEQ responses) indicating a good ability to predict the overall rating of the health states. The observed and predicted rating scores for the PEQ health states were largely similar; again indicating a good model fit Appendix 27). The co-efficients from the multinomial regression model were then applied to the ePAQ trial data to estimate the probability of membership into each rating score group based on combinations of participants PEQ responses.

The results from the application of the base case regression model (which included the process domain, severity and interaction and are reported in Table 28), to the ePAQ trial data are reported in Table 29. The analysis included all ePAQ trial patients who had complete cost, SF-6D and PEQ data, and had a rating score calculated for the communication domain based on their PEQ responses. This resulted in 26 patients from the intervention and 27 from the control group.

Descriptive analysis

Prior to replicating the best practice methods for analysis described in Chapter 2, a descriptive analysis is given so that the construction of the utility decrements is transparent. Once included within a statistical analysis, the impact of these various steps becomes more difficult to disentangle.

Table 29 shows the application of the process utility decrements to the ePAQ trial results. Mean total costs were calculated for each trial arm, from an NHS perspective. Costs were slightly higher in the intervention arm than the control (£15.17). Mean SF-6D values were calculated at baseline (t_1) and 6 months (t_2). This was used to calculate the QALY gains for the analysis without process.

In order to calculate the process utility decrement, the mean process level was calculated for each trial arm, from the application of the multinomial regression to the ePAQ data. This provided a process level for each patient based on their PEQ responses. The mean severity of the underlying health states for each arm was

calculated based on patient's baseline SF-6D (from the ePAQ trial) and the use of dummy variables (0 for sick patients, and 1 for healthy). The regression co-efficients (for process level and severity) were applied to these mean values to calculate a process utility decrement for each individual patient based on their individual PEQ results. This was used to adjust SF-6D t_2 values to incorporate process, and produce an SF-7D value, which represents a change from baseline. Therefore, the SF-6D (t_1) and (t_2) are based on UK tariff values applied to responses within the ePAQ EEACT. The SF-7D (t_2) is based on the SF-6D(t_2) minus the process utility decrement.

The number of QALYs gained based on the SF-6D values, and SF-7D values was then calculated. The use of the SF-7D reduced the QALY loss associated with the intervention group from 0.0097 to 0.0070. Both analyses resulted in negative incremental cost-effectiveness ratios.

The cost difference including process is £15.17, and the QALY loss -0.007. The analysis reported here compares this to the ePAQ EEACT results Chapter 3, which from the NHS perspective showed a cost difference of £67.94 (when comparing intervention to control) and a QALY loss of -.0095.

Table 29. Application of process utility to the ePAQ trial data.

	Intervention N = 26	Control N =27	Difference
Total costs	£1,169.45	£1,154.28	£15.17
SF-6D t_1	0.6445	0.6175	0.0270
SF-6D t_2	0.6334	0.6259	0.0075
Change in SF-6D	-0.0111	0.0084	-0.0195
Mean process level*	1.5	1.78	-
Mean severity*	0.5	0.52	-
Mean process decrement	-4.9505	-5.5088	0.5583
Mean process utility decrement	-0.0495	-0.0551	0.0056
SF-7Dt2	0.5839	0.5708	0.0131
QALY Gained SF-6D	-0.0056	0.0042	-0.0097
QALY Gained SF-7D	-0.0303	-0.0233	-0.0070
IC/QALY (without process)	-£1,550.90		
IC/QALY (with process)	-£2,180.09		

*Co-efficients

Statistical analysis

A further analysis was undertaken which mirrored the methods for the handling of missing data performed for the ePAQ EEQACT reported in Chapter 3 (Table 31) where QALYs were lower in the intervention (-.0026) and costs were slightly higher (£15.10). The analysis was performed from an NHS perspective, in line with the perspective for the willingness to pay threshold of £20,000/QALY. The analysis again followed the work of Faria et al. (2014) and used the mi command in Stata®, and seemingly unrelated

regression (SUR) to account for correlation between costs and QALYs (Faria et al. 2014; Gomes et al. 2012). Outputs from the SUR can be seen in Table 30, where QALYs and costs are slightly higher in the intervention (.0026 and £4.31 respectively).

Table 30. NHS Costs and SF-7D

	Group	Co-efficient	Std.err	t	P>	95% CI	
NHS Costs	Intervention[1]	4.31	148.60	0.03	0.98	-287.27	295.89
	Control[2]	704.71	100.61	7.00	0.00	507.40	902.03
SF-7D	Intervention[1]	.0026	.0064	0.40	0.687	-.0101	.0153
	Control[2]	.0028	.0043	5.32	0.000	.01435	.0312

Table 31. Cost and SF-6D

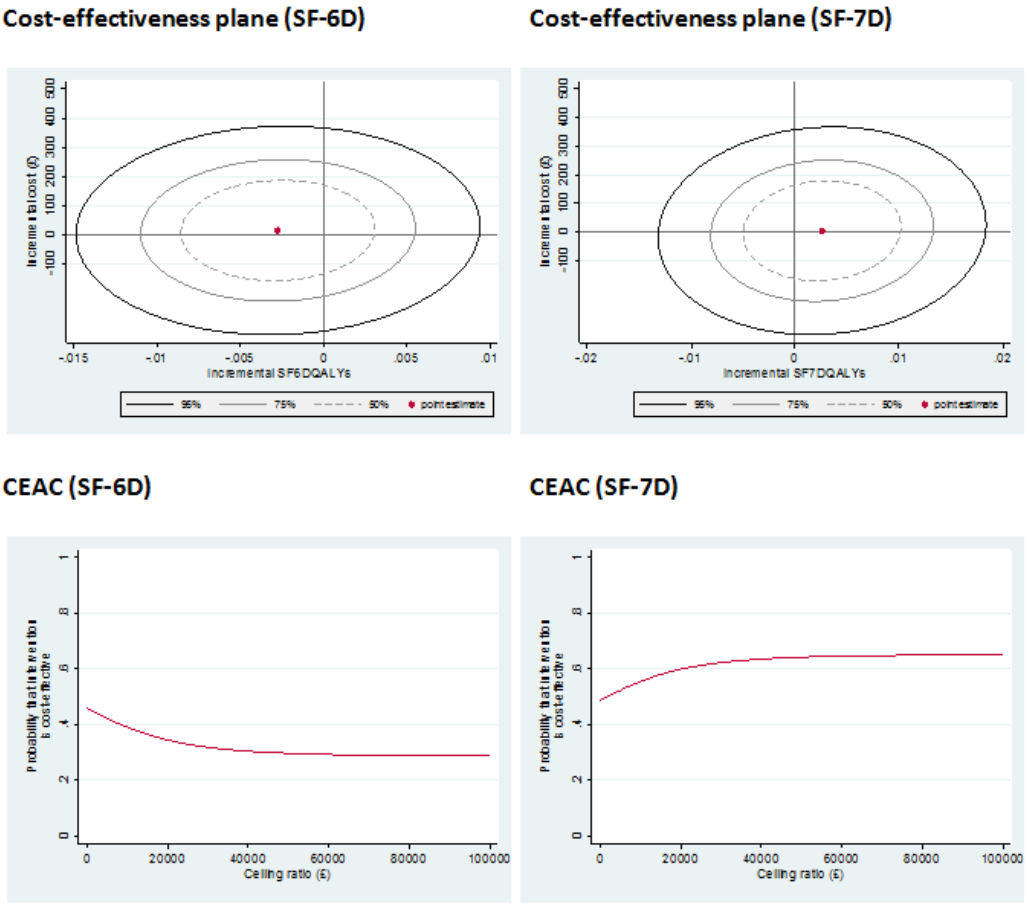
	Group	Co-efficient	Std.err	t	P>	95% CI	
NHS Costs	Intervention[1]	15.10	146.30	0.10	0.92	-271.87	302.08
	Control[2]	701.40	101.42	6.92	0.00	502.50	900.30
SF-6D[3]	Intervention[1]	-.00273	.00496	-0.55	0.58	-.01251	.00705
	Control[2]	.00259	.00354	0.73	0.47	-.00440	.00957

Figure 20 shows the SF-7D cost-effectiveness plane and cost-effectiveness acceptability curve alongside the SF-6D results which were estimated in Chapter 3. The SF-7D point estimate lays in the North East quadrant of the cost-effectiveness plane and the cost-

[1] Mean incremental value (compared to control)
[2] Mean value for the control group
[3] Change from baseline.

effectiveness acceptability curve (CEAC) shows there is approximately a 60% chance that the intervention is cost-effective at a threshold of £20,000 per QALY. The ICER is £1,656. (calculated from the difference in costs [4.31]reported in Table 30 divided by the difference in QALYs [.0026])

Figure 20. Cost-effectiveness planes, and CEACs for SF-6D and SF-7D analyses.



6.2 Discussion

This thesis suggests that in its current form, the QALY, which is the dominant measure of benefit within current economic analysis, does not capture all of the outcomes relevant to interventions, particularly where the expected outcomes of an intervention are non-health outcomes such as those associated with the processes of receiving care. This was illustrated by the ePAQ EEACT in Chapter 3. The trial, which compared the use of online ePAQ in advance of either a telephone (intervention) or a face to face (control) consultation, suggested a statistically significant difference between the control and intervention groups for three dimensions of the PEQ (communications, emotions and barriers), indicating that patients preferred the intervention. However, this preference for a telephone consultation was not picked up in the cost-effectiveness outcomes (SF-6D). This is shown in Figure 20 where the cost-effectiveness plane for the SF-6D analysis shows a slight additional cost and lower QALY in the intervention group. Consequentially, these central estimates sit in the north-western quadrant of the cost-effectiveness plane, where the intervention is dominated by the control, but with a 95% confidence ellipse that spreads to all four quadrants. There is approximately a 35% chance that the intervention is cost-effective at a threshold of £20,000 per QALY.

The SF7D analysis shows higher QALYs (0.003). Incremental costs change slightly, which may appear counterintuitive as the cost data remains unchanged from the SF6D analysis. However, this is because the correlation between QALY and costs has

changed, and this is included within the SUR estimation procedure. This does not change the results appreciably, and does not change the conclusions of this work in any way. These changes move the point estimates to the north-eastern quadrant of the cost-effectiveness plane, where the intervention is more effective and more costly than the control. The confidence ellipses also shift right, into the north-east quadrant and increase the probability that the intervention is cost-effective to around 60% (Figure 20). The ICER is £1,656.

Despite the low ICER, there remains a lot of uncertainty due to the small cost and effect sizes, as well as the impact of missing data. There is also considerable methodological uncertainty around the SF-7D estimates. One aspect of this uncertainty relates to the expected duration of the process utility gain. Within the current approach the EEACT has a 6 month time horizon, and the process utility gain is assumed to last for 6 months also. Whilst it could be argued that, for some, their communication experience with the doctor may affect them for 6 months following the consultation, others may feel it only applies to a shorter time frame. This question of how best to apply process utility needs further consideration, particularly in relation to the expected duration of the effect, and whether the process effect has the same duration as the effects of health outcomes. Although the literature reviews identified within the process utility literature review reported in Chapter 4 considered the timing of the measurement of non-health outcomes, only Brennan & Dixon (2013) highlighted this issue, but none specifically discuss the duration of the effects and its implications for economic evaluations (Ryan et al. 2014; Higgins et al. 2014; Opmeer et al. 2010).

Despite these uncertainties, the objective of this analysis was not to produce definitive cost-effectiveness recommendations; it was to demonstrate differential cost-effectiveness after the inclusion of exploratory process utility estimates. In essence, this work 'closes the loop' on the methodological work by highlighting it's applicability to an EEACT and its potential impact on cost-effectiveness estimates.

The estimated incremental costs and QALY in these analyses differ from those shown in the descriptive analysis in Table 29. The principal reason for this is the use of imputation of missing data, whilst the use of SUR is also expected to have had a small impact. However, for the purposes of this analysis, the various assumptions that are implicit within imputation and regression analysis are not considered important.

The results also display some face validity, with overall worse health states having lower utility values. The results also indicate a greater decrement with worse process, although it failed to show whether the decrements were greater in more or less severe patients, as was the case in the study by Brazier et al. (2010). When adding the process attribute to the SF-6D health states, the selection was based on the methodology by Krabbe et al. (1999) where levels of the cognitive domain which were added to the EQ-5D were chosen randomly, but avoiding a level which would be unlikely to have in combination with the other 6 levels. However, in retrospect, as the additional domain was a non-health domain, this step of the selection processes was not necessarily indicated, and resulted in some instances in more mild domain levels being added to the more mild SF-6D health states, and more severe domain levels

being added to the more severe SF-6D health states. This may have impacted on the results by over-emphasising the difference in valuations between mild and severe process domains. The findings do however support the findings by Brazier at al. (2010) whereby the addition of a domain, whether good or bad, lead to a reduction in the health state valuation in their study which added a pain dimension onto an asthma questionnaire. They suggested this may be due to a "focusing effect", however, the utilization of PEQ data already collected from with the ePAQ trial within our study aimed to reduce this focusing effect, and increase the internal validity, and there may therefore be alternative reasons for this trend which we are unable to identify within this study.

The findings from this study indicate that, just as others have added health domains to existing preference measures (Brazier et al. 2010; Swinburn et al. 2013; Yang et al. 2008), it is also possible to bolt-on non-health domains such as those related to process. However, just as Swinburn et al. (2013) suggest the development of bolt-on dimensions should include qualitative work to ensure the information being captured is the information which is important to patients, a qualitative approach would also be needed to fully understand respondents' reactions to these combined health states, and the implications of this for the valuations obtained.

There were some findings which were concerning. Although exploration of the mean differences between successive process domain levels showed overall greater differences for more severe domains, there were 74 respondents who provided

irrational responses to the level effects questions: where the addition of a worse process domain resulted in a higher valuation than when adding a better domain level. Again, further qualitative work into the valuation of these health and process states would be valuable.

This study has furthered the work of others. It builds on work by Krabbe et al. (1996) who also used the VAS, but in a study bolting on a cognition domain to the EQ-5D, by developing the additional domain from an existing, and psychometrically tested instrument (the PEQ). The PEQ is a valid measure of process which was used within the ePAQ trial and was utilised for the bolt-on study. Where Krabbe et al. (1996) used a convenience sample identified on a university campus, and Brazier et al. (2010) targeted one locality of the UK, this study used a market research company to obtain a sample which was broadly reflective of the general population. Cohort characteristics for the base-case population were largely comparable between groups, although more men than women completed the SF-6D (54% versus 48.5%), and more women than men completed the SF-7D (45.4% male, versus 51.5% female). There were also differences when comparing cohort employment status with national census data. The 2011 census reports that 13.10% of the population were retired. This compares to 45.4% (SF-6D) and 39.2% (SF-7D) for the study cohort, and may impact on the generalisability of the results.

There was also a relatively high drop-out rate of 66% of those who began the survey but dropped out before completion. The reasons for this seemingly high rate are not

known although, it does compare to other online health state valuation studies. For example, Bansback et al. (2014) reported a 46% drop-out rate for a discrete choice experiment study, and Mulhern et al. (2014) reported a 50% drop out rate for a study into EQ-5D-5L health state preferences.

Despite the generally positive findings from this study, it should be noted that at this stage the work is exploratory. Certain decisions were made to simplify or overcome methodological barriers in order to ascertain the feasibility of undertaking further, more detailed analysis. These include:

- The PEQ communication domain was used to represent process; it was selected partly because the ePAQ EEACT results indicated a significant difference in the PEQ communication results between study arms. We would therefore expect the differences in process utility which were detected.
- The PEQ domain was mapped across to a different scale that was developed to make the elicitation exercise easier for respondents. Pilot work suggested that simply transferring PEQ question into a health state description was potentially problematic, therefore an alternative approach was taken, this incorporated additional uncertainty relating to the mapping process and some methodological uncertainty relating to the validity of the new scale.
- A reduction in the number of process domains from 5 to 3. The descriptive analysis of the 5 process domains did suggest greater utility decrements with more severe process domains; however, there was a clear monotonic relation-

ship when reduced to 3. As the purpose of this study was to test the acceptability of the methodology, the decision was taken to use the 3 domain data.

- The severity of the health state was determined using dummy variables, 0 = sick patients, or 1 = healthy patients. It may be that a more sophisticated approach could be taken to indicate severity. As there were only 22 data points (health states valued), and only 260 respondents, it was decided that this simplistic representation of severity would explore the impact of the severity of the underlying health state, but not over-complicate the analysis, which, at best, was likely to show only minimal differences in utilities.

6.2.1 Recommendations for future research

This study suggests it is possible to capture process through a bolt-on study. However, given the exploratory nature of the study there are areas which require further work. The decision was made to reduce the number of process levels from 5 to 3. Further analyses could be undertaken to run regressions on the 5 level data, to explore the implications of the simplification by comparing with the 3 level analysis. This study assumed that the impact of the dimensions on preferences is additive, however, this assumption has been questioned by Brazier et al. (2010) and has not been addressed within this process bolt-on study. If further studies are developed which expand on the QALY to incorporate additional dimensions then this should be explored further, and the impact on the resulting utilities should be considered. The observation that the severity of the underlying health state impacts on the utility decrements derived

should also be explored further, as well as the expected duration of process utility gain and its application in economic evaluations.

In addition to this quantitative work, further qualitative work would be valuable. The number of irrational responses identified in this study highlighted a potential concern. Insight into respondents decision making processes, which could be gained through qualitative research may provide important insight into this underlying reasons for this, and allow us to understand more fully the utility estimates derived.

6.3 Conclusion

As a feasibility exercise, this quantitative study shows that it is possible to develop a process domain from an existing psychometrically tested measure, and it is also possible to bolt it onto an existing PBM. This is the first study which bolts on a dimension to the SF-6D. This is important, as there are more domains in the SF-6D, than the PBMs used in previous bolt-on studies, and the descriptors are more complex, for example they have 4 to 6 levels. It is therefore encouraging that the process utility values which were derived from this do appear to have some face validity, and there is the potential for their inclusion to change the ICER and CEAC.

Despite these findings however, there is uncertainty introduced by some of the methodological assumptions, which may have impacted on the study results. There is also the issue that this study was based on VAS valuations, when choice based valuations are preferred. It is not therefore known whether people will be willing to

trade-off health for different levels of process, or whether the questions will seem nonsensical. Chapter 7 will therefore take a qualitative approach to explore this further.

6.4 References

- Bansback, N. Hole, A. Mulhern, B. Tsuchiya, A. (2014). Testing a discrete choice experiment including duration to value health states for large descriptive systems: Addressing design and sampling issues. Social Science and Medicine, 114: 38-48

- Brazier, J. Rowen, D. Tsuchiya, A. Yang, Y. Young, T. (2010) What a pain: adding a generic dimension to a condition-specific preference-based measure. *HESG Abstract*.

- Brazier, JE. Roberts, J. (2004) The Estimation of a preference-based measure of health from the SF-12. *Medical Care,* 499, 851-859.

- Brennan, VK. and Dixon, S. (2013) Incorporating Process Utility into Quality Adjusted Life Years: A Systematic Review of Empirical Studies, Pharmacoeconomics, 31(8):677-91

- Faria, R. Gomes, M. Epstein, D. White, IR. (2014) A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics,* 32(12),1157-70.

- Gomes, M. Grieve, R. Nixon, R. Ng, ES. Carpenter, J. Thompson, SG. (2012) Methods for covariate adjustment in cost-effectiveness analysis that use cluster randomised trials. *Health Econ*. 21(9),1101-18.

- Higgins, A. Barnett, J. Meads, C. Singh, J. Longworth, L. (2014) Does Convenience Matter in Health Care Delivery? A Systematic Review of Convenience-Based Aspects of Process Utility. *Value in Health*, 17,877-887.

- Krabbe, P. Stouthard, M. Essink-Bot, M. Bonsel, G. (1999) The effect of adding a cognitive dimension to the EuroQol Multiattribute Health-Status Classification System. *J clin Epidemiol,* 52(4), 293-301.

- Mulhern, B. Bansback, N. Hole, A. Tsuchiya A. (2014) A methodological exploration of designing DCE with duration to model EQ-5D-5L health state preferences: The first paper from the FEDEV project. EuroQol Group Plenary, Stockholm.

- Opmeer, BC. de Borgie, CA Mol, BW. Bossuyt, PM. (2010) Assessing Preferences Regarding Healthcare Interventions that Involve Non-Health Outcomes: An Overview of Clinical Studies. *Patient*, 3(1), 1-10.

- Ryan, M. Kinghorn, P. Entwistle, VA. Francis, JJ. (2014) Valuing patients' experiences of healthcare processes: Towards broader applications of existing methods. *Social Science and Medicine* 106,194-203.

- Swinburn, P. Lloyd, A. Boye, KS. Edson-Heredia, E. Bowman, L. Janssen, B. (2013) Development of a disease-specific version of the EQ-5D-5L for use in patients suffering from psoriasis: lessons learned from a feasibility study in the UK. *Value in Health, 16,1156-1162.*

- Torrance, GW. (1986) Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5, 1-30.

- Yang, Y. Brazier, J. Tsuchiya, A. (2008) The effect of adding a 'sleep' dimension to EQ-5D. Health Economists' Study Group meeting January.

Chapter 7: Think Aloud Study Literature Review and Study

7.1 Introduction

The process utility valuation study indicated that it is possible to identify and measure process utility. However, this depends on the respondents' ability to value health scenarios which include both health and process outcomes. We should also therefore consider whether respondents find it sensible and meaningful to value combined health states, what personal and subjective factors they consider in the valuation process, and what this means for the health state values derived from the valuation study reported in Chapters 5 and 6. This was therefore explored further with a qualitative study.

There is a body of quantitative literature exploring different aspects of health state preference studies, for example, van Osch and Stigglebout (2005) interviewed respondents following completion of a VAS valuation to explore approaches to valuing health states, and Devlin et al. (2004) performed a qualitative analysis of respondents' written comments from a self-completed EQ-5D health state valuation questionnaire. However, to date there has been little associated qualitative work undertaken to gain in-depth insight into the human decision making processes associated with such studies. One qualitative method which has been used within this area of health economics is verbal protocol analysis (VPA). This is a method which is used to collect and analyse data on the cognitive processes associated with the completion of a task (Miller & Brewer, 2003). VPA involves making a detailed record of a person's verbal

report whilst they are carrying out a specific task (concurrent), or sometimes immediately following completion of a task (retrospective) (Ball et al. 1998). These verbal reports are developed by asking the participant to "think aloud" during the task completion (Ericcson & Simon, 1993). For this reason some also refer to this technique as "think aloud" (Miller & Brewer, 2003). Think aloud, rather than the term VPA will be used to describe the technique within this thesis.

Unlike traditional qualitative interviewing, the think aloud technique is specifically for use during task completion, this is ideal for this study which aims to understand people's thought processes as they value health scenarios combining health and process outcomes. Asking respondents to think aloud concurrently with completing the task reduces the time they have to consider their responses, and aims to reduce bias. In order to gain an understanding of the status of current health economics literature using the Think Aloud (TA) technique a targeted literature review was performed.

7.2 A Literature Review of Think Aloud studies.

7.2.1 Aims and objectives

The aims and objectives of this literature review were to:

- Provide an understanding of the extent of the current literature applying the TA technique to health economic principles;
- Gain an understanding of the methodologies applied;
- Inform the design of a think aloud study.

7.2.2 Literature Review Methodology

A focused literature search was undertaken. This approach was chosen through discussion with my supervisors, for several reasons: The focus of this thesis is process utility as opposed to qualitative research methods; it was felt that a systematic review was not necessary to identify good practice within the qualitative exploration of health state valuation studies, and that all of the relevant studies would be identified using a targeted approach. This approach was consolidated by comparing the included studies with a fellow PhD student also undertaking a Think Aloud study as part of his PhD thesis. If this qualitative research study had been central to this thesis, a systematic literature review would have been seen as essential.

Searches were performed in October 2013 and updated in July 2015. The searches were performed in the following databases:

- Pubmed
- Cochrane library (a methodological search using the terms “think aloud” and “verbal protocol”)
- Web of Science (formerly Web of Knowledge).

The following combinations of search terms were used:

- Think aloud AND economics;
- Verbal protocol AND economics.

No date or language limits were applied. All studies which applied a think aloud/verbal protocol analysis methodology within the health economics field were included in the review. The reference lists of all included studies were also checked for additional studies which met the inclusion criteria. The list of included studies was shared with a fellow PhD student who included a similar review into Think Aloud studies in health economics in his thesis.

7.2.3 Literature Review Results

The review identified 10 studies. An additional 2 studies were provided by a fellow student. A further paper was provided by the author, a member of staff at ScHARR who VB met with to discuss the study design (Mulhern et al. 2012). Table 31 provides an overview of all 13 studies.

Table 32. Summary of Think Aloud Studies in Health Economics.

Author (year of publication)	Aim	Elicitation method	Questionnaire	Sample size	Population	Quantitative	Qualitative: Semi-structured interview
Al-Janabi et al. (2013)	To investigate the feasibility of individuals reporting their own capabilities using the ICECAP-A and EQ-5D.	-	EQ-5D ICECAP-A	34	general population	No	Yes
Baker and Robinson (2004)	Qualitative study alongside a SG preference elicitation study to explore how respondents answer SG questions.	SG	N/A	31	characteristics of patients with HT	Yes utility estimates reported for 8 health states	Yes
Cheraghi-Sohi et al. (2007)	Think aloud study into a DCE evaluating patient priorities in primary care	DCE	N/A	20	University staff.	No	No

Author (year of publication)	Aim	Elicitation method	Questionnaire	Sample size	Population	Quantitative	Qualitative: Semi-structured interview
Damschroder et al. (2005)	To explore whether people mention considerations related to distributive principles rather than QALY-maximisation more often in PTO or TTO elicitations.	PTO TTO	N/A	64	General population	yes	yes
Horwood at al. (2014)	Evaluating the face validity of the ICECAP-O Capabilities measure: A "think aloud" study with hip and knee arthroplasty patients.	N/A	ICECAP-O	20	Patients with OA knee/hip	No	No
Mulhern et al. (2012)	Exploration of how respondents perceive health state valuation exercises using TTO and DCE.	TTO, DCE	N/A	30	Convenience sample	Yes	Yes
Osch and Stigglebout (2008)	To explore SG biases using qualitative data.	SG	N/A	45	General population	Yes	No
Papageorgiou et al. (2014)	To explore the feasibility of TTO to elicit valuations for depression experienced alongside a somatic illness.	TTO	N/A	10	Convenience sample	Yes	No

Author (year of publication)	Aim	Elicitation method	Questionnaire	Sample size	Population	Quantitative	Qualitative: Semi-structured interview
Robinson et al. (1997)	To gain insight into three key questions which emerged from the MVH study.	VAS TTO	N/A	45	Respondents of the MVH study	No	Yes
Ryan et al. (2009)	To investigate if respondents hold complete, monotonic and continuous preferences when completing a DCE.	DCE	N/A	18	General population	Yes	Yes
Spencer 2000	Testing the additive independence assumption in the QALY model using a SG	SG	N/A	29	General population	Yes	No
Spencer 2001	To explore the issues considered by respondents completing a TTO questionnaire to value EQ-5D.	SG	N/A	30	General population	Yes	No
Whitty 2014	To compare the validity and acceptability of discrete choice and best worse scaling methods.	DCE, BWS	N/A	24	University students and staff	No	No

BWS = Best worse scaling; DCE = discrete choice experiment; HT = hypertension; ICECAP-A = a measure of capability for the adult population; ICECAP-O = a measure of capability for older people; LT-TTO = lead-time time trade off; SG = standard gamble; OA = osteoarthritis; PTO = Person trade off; TTO = time-trade-off; VAS = Visual analogue scale.

Aims of included studies

All studies applied qualitative techniques to principles routed in health economics. In most cases the studies aimed to gain insight into the issues considered by respondents when using preference measurement techniques such as SG (Van Osch & Stigglebout, 2008) and TTO (Robinson et al. 1997; Papageorgiou et al, 2014) to value health state descriptions. For example, Baker and Robinson (2004) undertook a study which used SG to elicit values for health states associated with anti-hypertensive medication. In order to understand the thought processes respondents were going through in the completion of the SG tasks, respondents were asked to "think aloud" during the SG exercise, and to subsequently complete a semi-structured interview. Van Osch and Stiggelbout (2008) and Spencer (2000; 2001) also explored issues surrounding preference measurement using the SG technique.

Damschroder et al. (2005), Robinson et al. (1997), Papageorgiou et al. (2014) and Mulhern et al. 2012 all studied the TTO. Damschroder et al. (2005) used TA techniques to compare whether respondents found principles related to the distribution of resources more important in completing person-trade off (PTO) elicitations than in TTO elicitations. These 2 approaches differ in that TTO asks respondents to state their preference for a health state by imagining they are in that health state, or that they could be returned to full health but for a shorter number of years; PTO asks people to make trade-offs between treating different groups of patients. The authors suggest the latter are more likely to capture preferences in regards to not only the health states

themselves, but the distribution of health care resources to society. They therefore presented respondents with both a TTO elicitation and a PTO elicitation, and were asked to think aloud whilst completing them. Robinson et al. (1997) used qualitative methodology to explore questions which arose from the findings of the MVH study (which used TTO and VAS to determine a set of relative valuations for health states in the UK general population). The 3 key questions were: Why are some states that are rated better than dead on the VAS often rated as worse than dead in the TTO?; Why are some respondents unwilling to trade off any time at all in order to avoid a health state that they place below full health on the VAS?; and why are TTO valuations of older respondents for the more severe health states lower than those of the younger age group? To explore these questions, respondents were asked to complete questions taken from the MVH study whilst thinking aloud. Papageorgiou et al. (2014) used think aloud techniques to assess the feasibility of using TTO to value depression experienced alongside a somatic illness. Mulhern et al. (2012), used think aloud techniques to test 4 types of valuation tasks. These included variations of TTO and DCE. They aimed to investigate the acceptability of the exercises to lay respondents, the strategies used to complete the tasks and whether the issues impacted on the validity of the exercises. Their study also aimed to investigate specific issues relating to the EQ-5D-5L descriptive system, for example, how the severity levels were interpreted, and whether there was any variation according to the valuation method.

All the studies discussed so far investigate approaches which are commonly used to estimate utility values for health states, and whose development is rooted in expected

utility theory. This refers to a set of axioms reflecting rational behaviour under uncertainty. Ryan et al. (2009) used an existing discrete choice experiment (DCE) questionnaire, which was used in previous research to investigate patients' preference for bowel cancer screening and asked respondents to think aloud whilst completing it. They used this approach to identify contradictions to these axioms (irrational behaviour under uncertainty). DCE was also studied by Whitty et al. (2014), who explored both the validity and acceptability of DCE and best worst scaling (BWS) using the think aloud technique, and Cheraghi-Sohi et al. (2007) who explored the application of DCE methods in primary care.

The remaining two studies were not specific to utility elicitation. Al-Janabi et al. (2013) asked respondents to complete the ICECAP-A (a measure designed for individuals to self-report their capabilities across 5 dimensions of life), and the EQ-5D, and aimed to explore the feasibility of individuals self-reporting their own capabilities. Horwood et al. (2014) examined the face validity of the ICECAP-O measure which is a version specifically designed for older people. They aimed to identify the extent and nature of problems that people encountered on completing the measure. Both studies used think aloud interviews during the completion of the measures.

Sample size and populations

Sample sizes ranged from 10 (Papageorgiou et al. 2014) to 83 (Robinson et al. 1997). Studies used either a patient sample (Horwood et al. 2014; Baker and Robinson 2004), a general population sample (Ryan et al. 2009; Spencer 2000, 2001; Al-Janabi et al.

2013; Damschroder et al. 2005; Van Osch and Stigglebout, 2008; Whitty et al. 2014; Cheraghi-Sohi et al. (2007), or a convenience sample (Mulhern et al. 2012; Papageorgiou et al. 2014).

Papageorgiou et al. (2014 p.3) included 10 participants, recruited through "informal networks", this included handing out invitations to participate within the university psychology department. Ryan et al. (2009), who performed a think aloud study of a DCE questionnaire for bowel cancer screening, recruited 18 respondents aged between 49-69 years from two bowling clubs in Aberdeen. For every participant in the study, the bowling clubs received £20. This is the target age group for bowel cancer screening. The sample was purposively selected to include both men and women from across the age range. This refers to a non-probability sample, in which the sample is selected in a considered way, so that the population studied are relevant to the research question posed (Bryman, 2008).

Horwood et al. (2014), who studied the ICECAP-O questionnaire, which is used for patients with osteoarthritis, used a similar sample size. They recruited 20 patients with osteoarthritis of the knee or hip, through orthopaedic clinics in the UK, excluding those with dementia, or who were unable to give informed consent. Again, participants were purposively sampled. This was based on social class, education, age and gender to reflect patients receiving hip and knee replacement surgery. They continued to recruit participants until a point of data saturation was reached, whereby no new themes emerged from the qualitative analysis.

Baker and Robinson (2004) also applied criteria to select patients who were typical in terms of age and gender to those relevant to their study into the valuation of health states for hypertension. Thirty-one patients typical in age and gender to those diagnosed with hypertension, but not themselves suffering from hypertension, were randomly selected from GP (General practitioner) records.

Robinson et al. (1997) included only participants who had previously completed the MVH (Measurement and Valuation of Health) study. Of 83 respondents based in the Northeast of England who had taken part in the MVH study, and expressed a willingness to be re-interviewed, 45 successful interviews were completed.

Al-Janabi et al. (2013) firstly identified a pool of potential participants by sending an invitation and screening criteria out to 600 members of the general population. For those who expressed an interest in participating, plus an additional 24 individuals who had been identified as potential participants in a previous study, purposive sampling was used to select a sample based on age, gender, ethnicity, health and socio-economic status, with the aim of obtaining a sample of socio-economic diversity. This resulted in a sample of 34 participants.

Both Whitty et al. (2014) and Cheraghi-Sohi et al. (2007) recruited participants through university email systems. Whitty et al. (2014) included 24 respondent aged 17 years and older and reported reaching data saturation. Cheraghi-Sohi et al. (2007) recruited 20 respondents.

The remaining studies reported no detailed criteria in the selection of participants. Damschroder et al. (2005) recruited a convenience sample of 64 members of the general population. Van Osch and Stigglebout (2008) recruited 45 respondents via newspaper advertisements and pamphlets and paid respondents 22.50 Euros to participate. Spencer (2000) recruited 30 members of the general population, who received a payment of £10 for participating, and Spencer (2001) recruited 30 members of the general public and mature students beginning a course at the University of York, and paid respondents £15 to participate. Mulhern et al. (2012) used a convenience sample, of non-academic staff at the University of Sheffield, identified via the university email system, and who received a £5 voucher for participating.

Think Aloud Interviews

As dictated by the inclusion criteria, all studies used the think aloud technique for the collection of qualitative data. The depth of reporting on the methodology differed between studies. An overview of the reported think aloud methodology reported by each study is provided in Table 33.

The think aloud technique can either be used whilst completing the task (concurrently), or following completion of the task (retrospectively). The approach taken was not reported in all studies, however, the majority used concurrent methods (Al-Janabi et al. 2013; Whitty et al. 2014), and some combined with retrospective clarification of queries (Ryan et al. 2009).

Seven of the studies included practice or "warm up" tasks to familiarise participants with the concept of thinking aloud (Al-Janabi et al. 2013; Horwood et al. 2014, Van Osch and Stigglebout, 2008; Ryan et al. 2009; Cheraghi-Sohi et al. 2007; Whitty et al. 2014). Al-Janabi et al. (2013) firstly, asked respondents to count the number of windows in their home, thinking aloud as they did so and secondly to complete a single five point question on their general health or life satisfaction. Horwood et al. (2014) asked participants to begin with 3 practice questions, although the questions themselves are not reported. Cheraghi-Sohi et al. (2007) began by asking basic arithmetic questions, and those who struggled were asked a further question about choosing a holiday destination. Ryan et al. (2009) and Whitty et al. (2014) warmed up with a practice game of noughts and crosses; and Van Osch and Stigglebout (2008) included 2 practice tasks but did not report the details.

In 7 of the studies, respondents were prompted to think aloud if they became silent for a period of time (Al-Janabi et al. 2013; Cheraghi-Sohi et al. (2007); Horwood et al. 2014; Van Osch and Stigglebout, 2008; Papageorgiou et al. 2014; Ryan et al. 2009). Al-Janabi et al. (2013) reported only prompting if they were silent for 10 seconds or more. Ryan et al. (2009) and Cheraghi-Sohi et al. (2007) also gave respondents clarification on the task if requested, and Damschroder et al. (2005) used verbal probing to ask the respondent to expand on think aloud statements both during and after the interview. Mulhern et al. (2012) asked respondents who provided an answer but did not successfully verbalise their thoughts, to give the reasoning behind their responses after answering.

Semi-structured interviews

Seven studies used semi-structured interviews in combination with the think aloud interviews to enhance the depth of information retrieved (Al-Janabi et al. 2013; Baker and Robinson 2004; Horwood et al. 2014; Ryan et al. 2009; Damschroder et al. 2005. Mulhern et al. 2012; Papageorgiou et al. 2014). These provided the respondent the opportunity to ask any questions resulting from the think aloud task, and also allowed the researcher to ask for clarification or further information on specific points which may have arisen.

Software used.

Six of the nine studies reported the use of specific software for the qualitative analysis. These were:

- Atlas-ti (Al-Janabi et al. 2013);
- HyperResearch©, (Spencer 2000, Spencer 2001);
- MS Word, used by (Horwood at al. 2014);
- Nvivo, used by: (Baker and Robinson 2004; Whitty et al. 2014).

Table 33. Summary of Think Aloud Interview Methodologies

Author (year of publication)	Concurrent/retrospective	Warm up questions	Prompting and probing	Analysis method reported	Additional information
Al-Janabi et al. (2013)	Concurrent	Two warm up questions: 1. Count the number of windows in their home 2. Complete a 5 point question on their general health	Researcher silent throughout TA task. Only interrupted if silent for 10 seconds or more. Asked to keep thinking aloud.	No specific approach reported, although used coding of "errors and struggles"	Semi-structured interview used to follow up on any questions respondents had expressed during think aloud.
Baker and Robinson (2004)	Not reported	Not reported	Not reported	No specific method reported, although included familiarisation, sorting, ordering and indexing, building themes, mapping and linking themes and interpretation.	Semi-structured interview followed think aloud.
Cheraghi-Sohi et al. (2007)	Concurrent	1. Basic arithmetic 2. Choosing a holiday destination.	Prompted if silent for "any length of time" (p278). Researcher only provided prompts and responded to occasional requests for clarification.	No specific approach reported, although included use of an existing coding framework and the incorporation of new themes. They report that the analysis was "in line with conventional procedures". (p279).	Nil

Author (year of publication)	Concurrent/retrospective	Warm up questions	Prompting and probing	Analysis method reported	Additional information
Damschroder et al. (2005)	Concurrent and retrospective	Not reported	Verbal probing whereby researcher prompted respondent to expand on their TA statements and/or to provide retrospective reported of their thoughts to elicit more complete verbalisation	Termed their analysis as a descriptive analysis, where they identified themes, and coded transcripts according to the themes.	Nil
Horwood at al. (2014)	Not reported (although appears to be concurrent with retrospective follow-up questions to clarify).	Asked 3 warm up questions. Questions not reported.	Researcher did not interrupt respondent unless they were silent for a few seconds, when the researcher reminded them to think aloud.	Segmenting, application of the framework used by Al-Janabi et al. (2013) to identify errors and struggles. They then reported using thematic analysis to explore participants' problems with completing the measure.	Researcher sat out of line of sight of respondent when completing think aloud. After completion of the questionnaire, the researcher asked any questions which had arisen.
Mulhern et al. (2012)	Concurrent and retrospective	Two warm up questions: 1) Asked how many windows are there in your house? 2) Presented with a DCE question about a choice of holidays. (same one as used by Cheraghi-Sohi et al. 2007)	If the respondent was unable to verbalise their thoughts they were asked the reasoning behind their decision after providing an answer.	Framework analysis	Nil

Author (year of publication)	Concurrent/retrospective	Warm up questions	Prompting and probing	Analysis method reported	Additional information
Van Osch and Stigglebout (2008)	Not reported	Given 2 examples to practice thinking aloud before each task.	If the subject became quiet during a task additional instructions were given to think aloud	No specific approach reported, although included familiarisation with the data and sorting and indexing of themes.	Nil
Papageorgiou et al. (2014)	Concurrent	Not reported	Occasionally used motivational probes to encourage thinking aloud.	A 3 step approach: 1. Interviews transcribed, and familiarisation 2. Quotes were compared for similarities or differences. Coding scheme was revised. 3. Quotes recoded into final coding scheme.	Nil
Robinson et al. (1997)	Not reported	Not reported	Not reported	No specific approach reported - report that a "verbal protocol analysis" was performed on all taped material, and responses were classified into broad categories. No further detail was provided.	Nil
Ryan et al. (2009)	Combination of concurrent and	Practice game of noughts and crosses.	If the respondent was silent for a period of time they were reminded to keep thinking	A charting approach whereby a chart was produced to record	Respondents were asked a de-briefing question.

Author (year of publication)	Concurrent/retrospective	Warm up questions	Prompting and probing	Analysis method reported	Additional information
	retrospective (approach selected from pilot study).		aloud. If respondents asked for clarification further explanation was given to maintain trust.	participants' performance and summarise the comments made.	
Spencer (2000)	Not reported	Not reported	Not reported	Based on "grounded theory". Included categorisation, coding, grouping of codes to form categories.	The interviewer verbalised respondents' hand non-verbal cues e.g. eye movements.
Spencer (2001)	Not reported	Not reported	Not reported	Minimal information reported. Report interviews were transcribed and analyses using HyperResearch ® software.	The interviewer verbalised respondents' hand non-verbal cues e.g. eye movements.
Whitty (2014)	Concurrently	Noughts and crosses	If after the first and 2 subsequent questions, the participant had not verbalised their thought processes the interviewer prompted "If you could keep thinking aloud"	Analyses "thematically" by developing initial framework and updating as new themes emerged.	At the end respondents were asked which 2 approaches (DCE or BWS) they preferred and why.

BWS: best worst scaling; DCE: Discrete choice experiment.

Analysis and Reporting of Results

Think Aloud Interviews

The analysis approach and depth of reporting varied between studies. Within all studies interviews were recorded and transcribed. An overview of the analyses reported can be seen in Table 32.

Al-Janabi et al. (2013) undertook their study to explore the use of two questionnaires for self-reporting capabilities. The think aloud interview was divided into 11 segments (5 representing items on the ICECAP-A measure, and 6 representing items on the EQ-5D). A coding framework was then applied based on the identification and coding of ”errors or struggles” respondents had in answering the questionnaire. The results then provided a description of these errors and struggles, incorporating illustrative quotations from respondents. This was followed by a discussion on the thematic analysis of completion, which combines findings from both the think aloud and semi-structured interviews.

Horwood et al. (2014) employ a similar methodology to Al-Janabi et al. (2013) in the exploration of a related questionnaire. They segment the interview (5 segments representing the ICECAP-O attributes), and then apply the same coding framework as used by Al-Janabi et al. (2013). The results include a table summary of the type and frequency of problematic segments for the participants', verbatim quotations which illustrate specific problems, and a summary of findings according to the identified themes.

Baker and Robinson (2004) reported the use of a **constant comparison method** to identify themes and the coding framework. Their analysis involved: initial familiarisation with the data; sorting coding and indexing; building themes; mapping and linking themes; and the building and interpretation of explanations. The analysis from the think aloud and semi-structured interviews was combined. The findings were reported according to each theme, with quotations used as evidence.

Damschroder et al. (2005) report conducting a descriptive analysis of the think aloud interviews. They identified themes and then coded transcripts according to these themes. Their reporting of results includes descriptive summaries, combined with tables of themes and illustrative quotations.

Whitty et al. (2014) reported that their Think Aloud data was analysed **thematically**, whereby they developed an initial coding framework, which they expanded as more themes emerged from the data. They report their results separately for each task (DCE or BWS) and include some narrative followed by quotations.

Van Osch and Stigglebout (2008 p.32) coded interviews using a "reference point" which was coded if a "comparison relative to a point of view was formulated." They completed initial familiarisation with the qualitative data, and sorting and indexing of relevant themes. Their results are reported according to the identified themes, and include some illustrative quotations.

Robinson et al. (1997) carried out a **verbal protocol analysis** on all transcribed material, where by respondents were classified into broad groupings. The results

were reported according to the classifications, with the number of respondents in each category reported, and quotations used for illustration and clarification.

Ryan et al. (2009) used a **charting approach** for the analysis of their think aloud interviews. This involved producing a chart to record each participant's performance on the quantitative tests and to summarise the comments they made during the interview in relation to each axiom of interest. The results are reported under headings related to complete, monotonic and continuous preferences, and include a table summarising both the quantitative and qualitative evidence.

Spencer (2000) based their analysis of the qualitative data on **grounded theory**. This involves categorising and coding data, grouping of codes to form general categories under particular themes. Cases which appear to contradict the general rules were then identified. The results are then reported according to each SG question, with a discussion on the qualitative findings, and illustrative quotations. Spencer (2001 p.15) provided less detail on the analysis approach. And report only that "interviews were transcribed and analysed using the software package HyperResearch© that aids in the management and analysis of non-numerical data." This analysis produced 14 general codes and 133 specific codes. The results were reported according to the 5 main themes which were identified, they report both descriptions of the findings according to each theme, and provide a table of the number of respondents displaying information related to each theme.

Papageorgiou et al. (2014 p.5) reported using a 3 step qualitative analysis which was "based on the principles of grounded theory". This involved transcription and

familiarisation, the grouping of quotes into a coding frame which was then revised into a final coding scheme. The interpretation of the data was then analyses using this coding scheme.

Mulhern et al. (2012) used a "framework approach" for the analysis of data, although no details on this approach were provided within the study, VB met with Brendan Mulhern at ScHARR to discuss. The framework approach was developed by Ritchie & Spencer, (2003) specifically for use in qualitative research related to policy decisions. It uses a matrix to summarise data and organise the analysis. Pope et al. (2000) suggest this framework, which is more rigorous than is usually imposed on qualitative research, assists in a completing a timely analysis within the timeframes imposed by policy driven objectives. The Framework approach consists of 5 stages: familiarisation; identification of a thematic framework; indexing; charting and mapping; interpretation (Pope et al. 2000).

Cheraghi-Sohi et al. (2014) did not report a specific analysis approach, however, they developed an initial coding framework from existing literature, which they adapted as other issues emerged.

Semi-structured Interviews

Five of the studies included semi-structured interviews. Al-Janabi et al. (2013) used a **constant comparison method** to derive explanatory themes. This involved the researcher constantly comparing the information being coded under a certain category so that a theoretical elaboration of the category can begin to emerge (Bryman, 2008). The remaining 5 studies (Baker and Robinson 2004; Damschroder

et al. 2005 Robinson et al. 1997; Ryan et al. 2009, Mulhern et al. 2012) combined the analysis from the think aloud and semi-structured interviews.

Study findings

All studies found the use of qualitative methods within the health economics field added to the existing body of literature. For example, Al-Janabi et al. (2013) showed that the majority of participants were able to comprehend questions on their capabilities, and but that they showed varying understanding of the concept of capability itself. Horwood et al. (2014. p.667) concluded that the think aloud interviews were "a valuable technique for pretesting the face-validity, acceptability and content validity of the ICECAP-O capabilities measure."

Of those studies exploring preference value elicitation, Baker and Robinson (2004) found that although overall respondents thoroughly considered the SG task when eliciting values, some respondents incorporated information which the authors felt was inappropriate to the decisions. For example, some found a situation where they would be offered surgery which had a 100% probability of death implausible and therefore disregarded this option.

Damschroder et al. (2005) questioned whether SG, TTO and VAS are suited to capturing public preferences for the allocation of resources in an environment of scarcity, and suggested that the use of PTO would be more appropriate. They found that respondents did consider principles related to the distribution of resources more than QALY maximising principles in the completion of PTO tasks, for example, the consideration of the importance of equal access to health care.

The study by Mulhern et al. (2012) into valuations of the EQ-5D-5L using TTO and DCE (within and without duration attributes added) found that respondents used a range of strategies to complete the valuation tasks, for example, they compared health state dimensions and severity levels. Their study found that although some respondents considered all of the attributes included within the health states in their valuations, others did not. Some focused on those most relevant to them, or those which were most severe. They also found that the addition of a duration attribute to the DCE increased the complexity of the task.

Papageorgiou et al's. (2014) feasibility study into the use of TTO for the valuation of health states which combine depression with a somatic illness, found that people had problems in imagining themselves living in the hypothetical health states, for example, imagining themselves suffering from depression for their whole life, or that someone suffering from depression and cancer would reach the age of 80 years.

7.2.4 Literature Review Discussion

The review identified 13 studies which used the think aloud technique in the health economics field. These studies most commonly explored preference measurement techniques such as SG and TTO to value health state descriptions (Baker and Robinson 2004; Van Osch and Stiggelbout 2008; Spencer 2000, 2001; Mulhern et al. 2012; Papageorgiou et al. 2014). Two of the studies used Think Aloud to evaluate responses to specific questionnaires (Al-Janabi et al. 2013; Horwood et al. 2014). None of the included studies used the technique to explore respondents' reactions

to the actual health states themselves as is proposed for the process utility think aloud study.

The sample size of studies varied considerably and ranged from 10 (Papageorgiou et al. (2014) to 83 (Robinson et al. 1997). Only one study provided reaching theoretical saturation as justification for the sample size (Horwood et al. 2014), whereby repeated interviews bring no new themes, and therefore there is no benefit gained from continuing interviews (Bryman, 2012). This may have implications for the study findings, particularly for studies with low sample sizes. Unless theoretical saturation has been met we cannot be sure that all of the themes associated with the tasks have been captured, or that the topic has been thoroughly researched. Further studies may therefore be required to confirm the study findings.

Studies used either patient or general population samples (including convenience samples). Of those studies exploring utility elicitation methods, three reported using purposive sampling, (Baker & Robinson, 2004; Ryan et al. 2009; Spencer 2011). Baker & Robinson (2004) purposively selected patients with the age and gender characteristics of patients with hypertension, although not suffering from the condition itself (Baker & Robinson 2004). Ryan et al. (2009) selected participants who fell within the target age for bowel cancer in their study into a DCE task for the elicitation of preferences for bowel cancer screening. Spencer (2001) included only patients aged between 21 and 59 years. People aged 60 years or older were not included as it was felt they might imagine dying before the end of the 10 year duration used in their TTO scenarios. Although neither Baker &

Robinson (2004) nor Ryan et al. (2009) report their justification for the sample selection, it may be assumed it is to address the issue identified by Papageorgiou et al. (2014) where respondents have difficulty in imagining themselves in some hypothetical health states. This issue is further confirmed by Robinson et al. (1997) who found that when using the TTO older respondents appeared less prepared to live for the next 10 years in reduced health states. Therefore, the age of the respondents is likely to impact on valuations. Although purposive sampling seems a sensible approach to take to address some of these issues in qualitative research studies, for the actual health state valuation studies which use TTO this is not strictly in line with the NICE recommendations, which recommend eliciting utilities which are reflective of a the general population (NICE, 2013).

Several of the non-patient samples were convenience samples, identified on university campuses (Whitty et al. 2014; Papageorgiou et al. 2014; Mulhern et al. 2012). When reviewing the findings from their studies, it should be remembered that they are gained from a sample that are likely to have higher education levels than the general population and may also be younger. The findings may not therefore be reflective of the population as a whole. The only study which used a patient sample was an exploration into the validity of a measure which is intended to be used in elderly patients, therefore justifying their sample selection (Horwood et al. 2014).

The depth of reporting and details provided on the think aloud methods themselves also varied between studies. Within think aloud studies, respondents can either be asked to verbalise their thoughts during the completion of a task (concurrent) or

after task completion (retrospective) (Kuusela and Paul, 2000). Seven of the included studies reported the use of concurrent/retrospective interviews (Cheraghi-Sohi et al. 2007; Papageorgiou et al. 2014; Ryan et al. 2009; Damschroder et al. 2005; Al-Janabi et al. 2013; Mulhern et al. 2012; Whitty et al. 2014). Ryan et al. (2009) explored 3 potential approaches within a pilot study: 1) Respondents were asked to verbalise their thoughts during completion of the questionnaire, but to expect interruptions; 2) Respondents were asked to think aloud when completing the task, with no interruption from the interviewers (except to be reminded to think aloud); 3) Respondents were asked to think aloud as they completed the DCS. If they were not thinking aloud then, after every second or third question, the interviewer asked them to reflect back on their choices. They found that this approach, which combines concurrent and retrospective methods, interfered less with the respondents thought processes, whilst still allowing further exploration where needed. Al-Janabi et al. (2013) chose a concurrent approach based on findings from an empirical study by Kuusela and Paul (2000) which compared verbal protocols collected concurrently, to those collected retrospectively, and found that the concurrent approach provided more fruitful information. As previously stated, Mulhern et al. (2012) asked some respondents to elaborate on their thoughts after valuing a health state, if they had failed to think aloud concurrently. Miller & Brewer (2003) suggest concurrent think aloud methods reduce the time respondents have to think about their responses and therefore reduces bias, this should therefore be considered when reviewing the study findings. Responses provided retrospectively may not provide a true reflection of respondents' thoughts during the completion of the task.

The analysis methods differed between studies. Two studies applied a "survey response model" which explores the "errors and struggles" experienced by respondents when completing surveys (Al-Janabi et al. 2013; Horwood et al. 2014). Damschroder et al. (2015) termed their analysis a "descriptive analysis" where they identified themes, and coded transcripts according to those themes. Ryan et al. (2009) used a "charting approach", whereby a chart is created for each participant recording the quantitative results, and summarising the comments collected from the think aloud interviews. This "charting" approach is not specifically referenced in the paper, although may be comparable to the "charting" stage of the framework approach developed by Richie & Lewis (2003). Papageorgiou et al. (2014) described their 3 step approach to the analysis where initially interviews were transcribed and a coding framework was applied, within the second step quotations were compared for similarities or differences and the coding scheme was revised. The quotations were then recoded into the final coding scheme. Mulhern et al. (2012) used the framework analysis approach, however, Ritchie & Lewis (2003) suggest that the depth of analysis using this approach remains dependent on the underlying study aims, therefore, all 5 stages may not be completed. This suggests the need to report not only the analysis approach, but a more detailed overview of the steps completed.

The remaining studies did not report specific approaches to their analysis of the data. However, overall, the majority of the studies did describe the different steps they applied. For example, Baker and Robinson (2004) reported undertaking familiarisation, sorting, ordering and indexing, building themes, mapping and

linking themes and interpretation. Although two studies reported only minimal information: Damschroder et al. (2005) identified themes and coded transcripts according to these themes, and Whitty et al. (2014) developed an initial framework which was updated as new themes emerged.

Six of the included studies performed semi-structured interviews alongside the think aloud interviews. The authors used these to ask for further clarification where needed on points raised by the think aloud interview, it allowed specific questions to be asked, and also allowed the respondent the opportunity to ask questions. The analysis and reporting of these interviews was less detailed than the think aloud interviews. Although A-Janabi et al. (2013) reported separate results and analyses for the information obtained through the think aloud and semi-structured interviews, the remaining 4 studies combined these analyses. If applying the comment from Miller & Brewer (2003) that information gained retrospectively may introduce bias this may be seen as a negative, however, it may also be seen as an enrichment of the information gained from the study.

Overall, the analysis methods seemed to vary considerably between studies, as did the depth of reporting. Following this review, there therefore remains a lack of transparency in the application of these qualitative research methods in health economics.

7.2.5 Literature Review Conclusion

This literature review identified studies which utilised think aloud methodologies in health economics. These studies were predominantly concerned with exploring

issues surrounding health state valuation tasks and were effective in generating qualitative data which, when analysed, provided new and useful insights into health state valuation tasks.

The study identified that clear reporting of the study methodology and analysis is vital to produce study findings that can be applied to clinical decision making. For example, it is important to know exact details on the study sample and justification for the sampling approach used. In order for us to be able to apply the findings to specific areas of interest detailed descriptions of the analysis process should be reported. This was evident in some studies and clearly lacking in others.

This review has been essential in informing a think aloud study into the valuation of health states including both health and process outcomes. It has confirmed that there is a body of literature which successfully uses think aloud techniques to explore utility elicitation studies qualitatively. These are of varying quality when considering aspects such as the study sample (true general population sample versus convenience samples), the use of concurrent versus retrospective think aloud techniques, and the necessity for the detailed reporting of the analysis approach.

The remainder of this chapter uses the information gained from this review to design and undertaken a think aloud study into the valuation of health scenarios combining both health and process outcomes.

7.3 A Think Aloud study

7.3.1 Think Aloud Aims and Objectives

The aim of this study was to apply the think aloud approach to methodology used in the process utility valuation study, to determine respondents' reactions to health states which combine both health and process outcomes.

It aimed to consider the following questions:

- Is it sensible to combine health and process outcomes within a health state?
- Can we accept the valuations derived from the valuation of these health states, for example those resulting from the process utility study reported in Chapters 5 and 6?

7.3.2 Think Aloud Methodology.

The methods were informed by the literature review into think aloud studies in Health Economics, from personal communication with staff/students within ScHARR who have undertaken think aloud studies, guidance from my supervisors, and the scientific review of the study protocol by Brendan Mulhern. Ethics approval for the study was provided by ScHARR Ethics Committee at Sheffield University.

Sample and Sampling

This study aimed to explore the thought processes respondents who completed the process utility valuation study (Chapters 5 and 6) underwent in deriving their responses. As preferred for the elicitation of utilities, the valuation study used a

general population sample, and therefore a general population sample was chosen for this study also.

Participants were recruited through the university internal emailing system. An email was sent to all students and staff at Sheffield University excluding those registered with ScHARR or the Economics Department, as it was thought they may have had previous knowledge and experience of health state preference studies which may contaminate the study findings. The email contained a brief introduction, and a link to a more detailed overview. Interested parties were asked to contact VB by email for further details (Appendix 28). Although this approach may have resulted in a convenience sample, as detailed in previous studies (Mulhern et al. 2012), this study used the purposive approach taken by Horwood et al. (2014). An initial pool of potential participants was identified consisting of all of the people who had responded to the original email, and as the number of responses allowed, respondents were asked for further details on their age and gender. This age and gender information was used to purposively select a sample of the general population based on Census age and gender data. All participants were aged 18 years or older. Selected participants were contacted by email, to arrange a suitable interview appointment. Those not selected were sent a response thanking them for their interest. A £10 gift voucher was be given to all those who completed the interview.

Valuation Method

The VAS was used within the process utility valuation study (Chapters 5 and 6). The aim of this think aloud study was to gain greater insight into the process utility

results. However, if process utility was used more widely, the valuation would need to be based on a choice based method (TTO or SG), which is the preferred method of utility elicitation (NICE, 2013; Brazier et al. 1999). These methods represent a different cognitive challenge to participants, therefore, respondents were asked to value the health states using both VAS and TTO techniques.

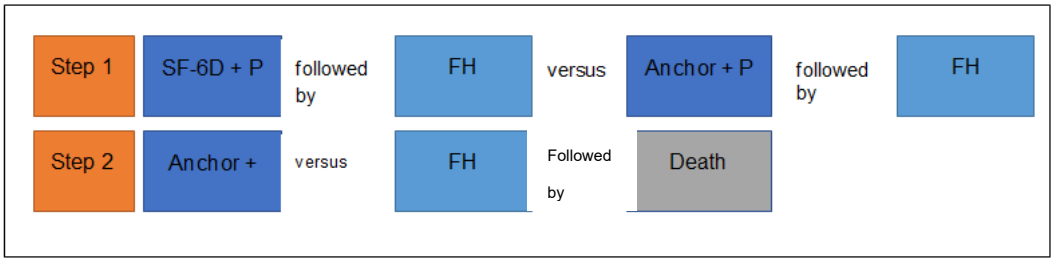
To summarise, the VAS uses a rating scale from 0 to 100, where 0 is the "worst imaginable health state" and 100 is the "best imaginable health state". Respondents are asked to locate the health state descriptions on the scale between the two anchors so that the intervals between the scores correspond to the differences they perceive (Drummond et al. 2005).

The TTO was developed by Torrance (1986) specifically for use in healthcare. The TTO questions were based on the MVH study (Dolan, 1997), which has provided the national values for the UK and has served as a guide for other national studies (Shaw et al. 2005; Lamers et al. 2006). It presents the respondent with a choice between two alternatives of certainty. When used to value chronic health states, respondents are asked to choose between a fixed duration of time in imperfect health, and a variable amount of time in perfect health, both of which are followed by immediate death. However, when used to value temporary health states, the movement from the health state to immediate death is thought to be unrealistic, and therefore an alternative "chained" approach is used. Within this scenario, the temporary health state is not compared to perfect health, or followed by immediate death, but is compared to a more severe anchor state, and followed by perfect health. This approach consists of 2 steps. Firstly, the temporary health state

is compared to the anchor health state, both of which are followed by full health. The second step involves comparing the severe anchor health state with normal health and death to transform it to a scale ranging from death (0) to full health (1) (Torrance, 1976).

The health states valued within this think aloud study were temporary health states, and it was felt that presenting the health states followed by immediate death would both be an unrealistic scenario, and may also inappropriately influence the respondents' thought processes. Therefore the two-stage TTO approach for the valuation of temporary health states was used to value the SF-6D and SF-6D+P health state descriptions. This can be seen in Figure 21.

Figure 21. TTO for the valuation of temporary health state i.



FH = Full Health; $_i$ = health state of interest; $_j$ = anchor health state (a state worse than $_i$); P = process domain.

In step 1, respondents are offered 2 alternatives:

- The health state of interest (SF-6D+Pi) for a fixed time (*t*), followed by full health versus,
- An anchor state (*j*) (worse than state i), for time (*x*) < *t*, followed by full health.

Time *x* is varied until the respondent is indifferent between both health states. At this point the preference value for state j is $h_i = 1-(1-h_j)\ x/t$. Within this step, *t* was set at 4 weeks on the assumption that this represented a plausible length of time that a bad experience of a health care contact may persist with a patient. As it was deemed unlikely that a patient would attend a consultation with the doctor when in perfect health, the P-domain was not combined with the perfect health state.

To transform the score onto a scale anchored by 0 (dead) and 1 (full health), as required for utilities, the anchor health state j was re-defined as a short duration chronic health state. In step 2, respondents were offered 2 alternatives:

- The anchor health *j* state for time *t*, followed by immediate death.
- A fixed time (*x*) in full health, where $x<t$, followed by immediate death.

The value of the anchor health state j = $x/(x-t)$. This value was applied in the step 1 formula [$hi = 1-(1-h_j)\ x/t$] to obtain a value for h_i.

Within this step, *t* was set at 5 years. Whilst this may be unrealistic for a health state that includes a process domain, it was hoped the think aloud study would also provide insight into whether a 5 year period is found to be implausible.

Pilot Study.

A methodology was initially developed and tested through a pilot study. The pilot study consisted of a convenience sample of 4 participants. One was male and 3 were female, and the ages ranged from 35 to 68 years. The pilot interviews were performed by one researcher (VB), and held at the participants own homes.

Health States

Health states were selected by the lead researcher (VB) and were initially chosen to mirror patients included within a clinical trial which was used to develop the process utility valuation study. The following health states were therefore selected:

- 111112(1) [Most commonly experienced health state in ePAQ EEACT].
- 121122(4)
- Anchor health state: 243223

However, after completing the 4 pilot interviews, it was evident that the respondents found the health states so mild that they were unwilling to trade a time of 1 week or more of life. Therefore, the following, more severe health states were selected:

- 343423(4)
- 343323(2)
- Anchor health state: 345554

Values for two SF-6D and two SF-7D health states were derived. Step 1 was completed for all 4 health states, and step 2 was performed twice (for SF-6Dj anchor state, and SF-6D+Pj anchor state). In total respondents completed 5 VAS, and 6 TTO valuations. As the pilot interviews lasted for between 52 minutes and 1 hour 5 minutes, it was thought that additional health states would excessively increase the burden of completion.

Additional pilot study findings.

In addition to the changes to the health state selection, the pilot study provided further useful information:

- Respondents found the format of the TTO questions difficult to understand. Therefore the wording in the questions, and the table for response included in the questionnaire booklet was adapted for greater clarity.
- The pilot study also provided useful information for the timing of interviews. Based on the duration of the pilot interviews (between 52 minutes and 1 hour 5 minutes), 1.5 hours were allowed for each study interview.

Interview Technique.

All study interviews were performed by one researcher (VB) and held in a private room at Sheffield University between May and November of 2014. Respondents were initially provided with a copy of the study information leaflet they had previously been sent via email. They were then asked to read and sign a consent form, which included permission to record the interviews. The interview consisted of two components: A think aloud interview followed by a semi-structured interview.

Think Aloud Interview.

The researcher explained the study and obtained informed consent. An initial warm up exercise was undertaken whereby respondents were asked to think aloud when responding to the question "how many windows are there in your house?" This approach was used in previous studies (Al-Janabi et al. 2013; Horwood et al. 2014,

Van Osch and Stigglebout, 2008; Ryan et al. 2009), and is recommended by Ericcson and Simon (1993). The initial instruction and warm up question are reported in Figure 22. Respondents who found this difficult were asked a further question "Describe your kitchen to me". This second practice question was included at the recommendation of the Sheffield University Research Ethics Committee.

Participants were then asked to complete the survey which contained the VAS and TTO health state valuation questions. The interviewer (VB) remained in the room, sitting slightly behind the respondent to avoid distraction as recommended by Ericsson & Simon (1993), with the aim of reducing the interactions between the interviewer and participants. Participants were instructed to think aloud whilst completing the survey. This allowed the interviewer to record the thought process of the respondent concurrently as the tasks were being solved, rather than asking the participant to retrospectively remember what they did. If the respondent became quiet the interviewer provided a prompt "can you tell me what you are thinking?" The study used both concurrent and retrospective techniques. Whereby, if the respondent could not successfully verbalise their thoughts they were asked the reasoning behind their decision after providing an answer. If respondents had difficulty in understanding how to complete the task, for example, the TTO for health states worse than dead, VB gave minor advice, but ensured that this was not related specifically to the health states themselves.

Figure 22. Think Aloud instruction and warm up.

In this study we are interested in what you think about when you find answers to some questions that I am going to ask you to answer. In order to do this I am going to ask you to THINK ALOUD as you work through a questionnaire.

What I mean by think aloud is that you do not plan what you are going to say, but act as though you are alone in the room thinking out loud to yourself. Just say out loud whatever is going through your mind as you answer my questions, even if it seems obvious.

There is no right or wrong answer; we just want to hear how you think about these issues. It is most important that you keep talking. If you are silent for any long period of time I will remind you to keep talking out loud. Is there anything you would like me to go through again with you?

Good, now we will begin with a practice question to get you used to thinking aloud. I will ask you a question and would like you to tell me what you are thinking as you get to your answer: "How many windows are there in your house?"

(If respondents were able to Think Aloud no further warm up question was asked. For those who struggled with thinking aloud were asked a second warm-up question: "Please describe your kitchen to me.")

The order in which the health states were given differed between respondents:

- Half of the respondents were shown pairs of health states (SF-6D and corresponding SF-7D), and half were shown all of the SF-6D health states followed by all of the SF-7D health states.

- For half of the respondents the health states were ordered from mild to severe, and for the remaining half, health states were ordered from severe to mild.

These two different ordering approaches aimed to remove ordering effects as a potential confounder, thereby ensuring that any conclusions drawn were not wholly or partly due to the order used. Within quantitative studies this can be corrected for statistically. However, in qualitative studies it cannot.

Semi-Structured Interview.

The semi-structured interview immediately followed the think aloud exercise, and was performed by the same interviewer. The interview contained both open and closed questions and was used to:

- Explore the thoughts participants had whilst valuing the health states. For example, if they queried the relevance of a question, or appeared to struggle to answer it, the issue was explored in the semi-structured interview.
- Examine methodological research questions that have been highlighted.

A topic guide (Figure 23) was used to explore further respondents thoughts on the health states valuation task and to provide an initial exploration of 2 methodological issues with using the think aloud technique which have been highlighted previously:

1) Whether the think-aloud activity affects the performance of the main task;

2) Whether the presence of the researcher during the think-aloud activity influences what the respondent says (McGavock and Traeharne, 2011).

The topic guide was developed using information reported by Baker & Robinson (2004), and feedback from PhD supervisors (GJ and SD), and scientific review (BM). The topic guide was used purely as a guide, and VB used her discretion to adapt as appropriate for each respondent.

Figure 23. Topic Guide

Thank you. We have now finished the think aloud exercise. I am interested to hear some more of your thoughts on this.

- How did you find the think aloud exercise?
- Were there any questions that you found more difficult to answer than others?
  - If so, can you explain further to me how the question made you feel?
- Can you look back at question x, this question seemed to take you longer to complete than others can you tell me how this question made you feel?
- Were things important to you during the exercise?
  - If so, can you explain what these things were?
- What did you think about the health states themselves?
  - If answers were similar for SF-6D and SF-7D health states to probe for reasons.
    - Your answers for these 2 health states were very similar, can you explain more what you were thinking when valuing them?
  - If answers were different for SF-6D and SF-7D health states to probe for reasons.
    - Your answers for these 2 health states were very different, can you explain more what you were thinking when valuing them?
  - What do you think about including information on a consultation experience in the health state?
- Did the VAS and TTO make you think about things differently?
- Did this alter when communication was added to the health state description
- Was it easier to think of communication when trading off life, or harder?
- Did the 2 types of question within the TTO valuation make you think about things differently?
- Which task was the most difficult and why? Refer to the questions in the booklet and show:
  - VAS
  - TTO stage 1
  - TTO stage 2
- Do you think your responses to the questionnaire would have been the same or different if you were not required to think aloud when answering the questions?
  - Prompt: Would you think about things in the same way if you didn't have to talk aloud?
  - Prompt: Do you think your responses to the questionnaire would have been the same if I had not been in the room during the think-aloud activity?
- Is there anything else you would like to tell me about your experience of doing the think aloud exercise today?

Qualitative Data Analysis.

All interviews were tape recorded and transcribed. Data from both the think aloud and semi-structured interviews were analysed together. Five of the studies identified by the review also combined the analysis for the think aloud interviews with the semi-structured interviews (Baker and Robinson 2004; Damschroder et al. 2005 Robinson et al. 1997; Ryan et al. 2009; Mulhern et al. 2012). It was felt this approach would aid in the ease of analysis, and provide a more enriched outcome. Analysis was therefore based on the demographic questionnaire booklets completed by respondents when attending the interviews, memo's written throughout the data collection phase, and interview transcripts.

The analysis approach within qualitative research varies according to the aims and objectives of the study. The literature review also indicated that similar studies can be analysed using different approaches, dependent on the approach chosen by the research team. This indicates that the approach to the analysis of qualitative research can be chosen more freely by the researcher than in quantitative research. Given this, these data were analysed using the Framework approach developed by Richie & Lewis (2003), and previously used by Mulhern et al. (2012). There were 2 key reasons for this. Firstly, on comparing the analyses completed by studies included in the think aloud literature review, I preferred the structure imposed by this method. And secondly, Framework analysis results in a matrix consisting of rows of respondents' data broken down into the themes identified. This is included in the appendix of this thesis, and provides a transparent overview of the data for readers.

Pope et al. (2000) describe the 5 stages which can be included in a Framework analysis: familiarisation, identification, indexing, charting and mapping. All 5 of these steps were performed in the analysis. The **familiarisation** process involved revisiting the research proposal and study objectives, and existing papers in the field. It also included the transcription of all interviews. This was completed by VB, and all transcripts were imported into Nvivo 10 qualitative software.

In the **identification** stage, a thematic framework was developed. At the beginning of the identification state, two transcripts were reviewed by VB and a list of themes was generated. The initial framework was then independently reviewed by another member of the study team (GJ), comments were addressed and incorporated into an updated framework. This revised framework was then applied to 2 further transcripts, updated and re-reviewed by GJ. Following this iteration the framework was applied to all 20 transcripts – with some further refinement as new concepts and themes emerged.

The list was then used to construct an index, where linked concepts were grouped into themes, and sorted so there was an index with a hierarchy of main and subthemes. The researcher ensured there was an "other" category in each subset, to provide an identifier for any uncovered issues that arose from the broader area of concern. The thematic framework then had **index** numbers applied, using numeric codes, and was incorporated into the analysis.

Once the coding framework had been applied to all transcripts, a final review of the framework was completed by GJ on a sample of 2 randomly selected transcripts. All

transcripts were then reviewed once more to ensure all concepts had been identified and categorised according to the final framework structure. Respondent characteristics were also reviewed.

The coding framework which was developed can be seen in Figure 24. The coded data were then synthesised and sorted into a **thematic chart** in the fourth stage of the analysis. Descriptions of the themes were combined in a matrix with quotes from the interviews. The complete matrix charts can be seen in Appendix 32. Quotations were selected for inclusion into the main body of the thesis, with care being taken to ensure those selected represented a wide range of individuals and views.

In the **Interpretation,** the results for each theme which emerged were reported in detail, combining diagrams of key points, and selected quotations. For the purpose of reporting, study participants were assigned an identification code, where the initial letter (M or F) denotes respondents' gender.

Think Aloud Results

The recruitment email sent through the university internal emailing system (Appendix 31) resulted in a pool of 134 potential participants. An initial sample of 20 participants was selected based on an age and gender, to be largely representative of the general population. At this point the level of theoretical saturation was assessed, and it was felt that a point of theoretical saturation had been reached as no new themes were emerging. This approach was based on the Horwood et al. (2014) study, and a similar study performed by a PhD student in

ScHARR (Personal communication, PhD Student, Milad Karimi, 2013). These studies included 20 and 21 participants respectively before reaching theoretical saturation.

Respondent characteristics can be seen in Table 34. Just over half of respondents were male (55%), the mean age of all respondents was 30.8 years. Fifty percent of participants were employed, and over half of the participants were educated to degree level. On average the interviews lasted 43.1 minutes (range 25.9 – 54.7 minutes).

Nine main themes were identified. Eight of these related to the valuation of the health and process states, and one related to the impact of the think aloud technique itself. The coding framework can be seen in Figure 24.

Each of the themes is reported in turn, and flow diagrams are used to illustrate each further. Included in the flow diagrams is the number of respondents making comments related to each theme. These are reported purely to provide an indication of the number of times a theme was referred to, as opposed to reflecting the importance or significance of each theme. This approach is not usual for qualitative research; some suggest that "expressing results in relative frequencies may be misleading." (Pope et al. 2000, p.114). However, some find that simple counts provide a useful summary (Pope et al. 2000). As the framework approach facilitates this approach (a simple count of the number of quotations within each indexed theme), this information was reported.

Table 34. Respondent Characteristics.

Characteristic	N (%)
N	20
Gender	
Male	11 (55)
Female	9 (45)
Age in years	
Mean (SD)	30.8 (11.7)
Range	19-66
Marital Status	
Married/Partner	5 (25)
Single	15 (75)
Highest Education	
Education post minimum	8 (40)
Education to degree level	12 (60)
Employment status	
Student	9 (45)
Employed	10 (50)
Retired	1(5)

Figure 24. Thematic Framework

1. Interpretation of process domain
- 1.1. Do not understand
- 1.2. Interpretation of bad process domain
- 1.3. Interpretation of good process domain
- 1.4. Interpretation of process domain generally (not good or bad domain)
- 1.5. Other

2. Impact of having a consultation
- 2.1. Impact of having a positive consultation
- 2.2. Impact of having a negative consultation
- 2.3. Other

3. Scoring of health states (VAS)
- 3.1. Scoring of with process VAS health states
- 3.2. Scoring of without process health states
- 3.3. Other

4. Impact on preference for health states (TTO)
- 4.1. Process domain has no impact
- 4.2. Process domain influences decision
- 4.3. Other

5. Valuation method
- 5.1. Impact of time spent in health state
- 5.2. Ease of completion
- 5.3. Different thinking processes
- 5.4. Other issues on valuation method

6. Importance of process domain
- 6.1. Relative importance of process domain
- 6.2. Importance of process domain
- 6.3. Other information on importance of process domain

7. View on combining health and process
- 7.1. Positive view
- 7.2. Negative view
- 7.3. Relevant for some
- 7.4. Able to distinguish between health and process outcomes
- 7.5. Other

8. Heuristics
- 8.1. Simplifying heuristics

9. Think aloud interview

9.1. Views on technique

9.2. Impact of interviewer listening

9.3. Impact of TA on response

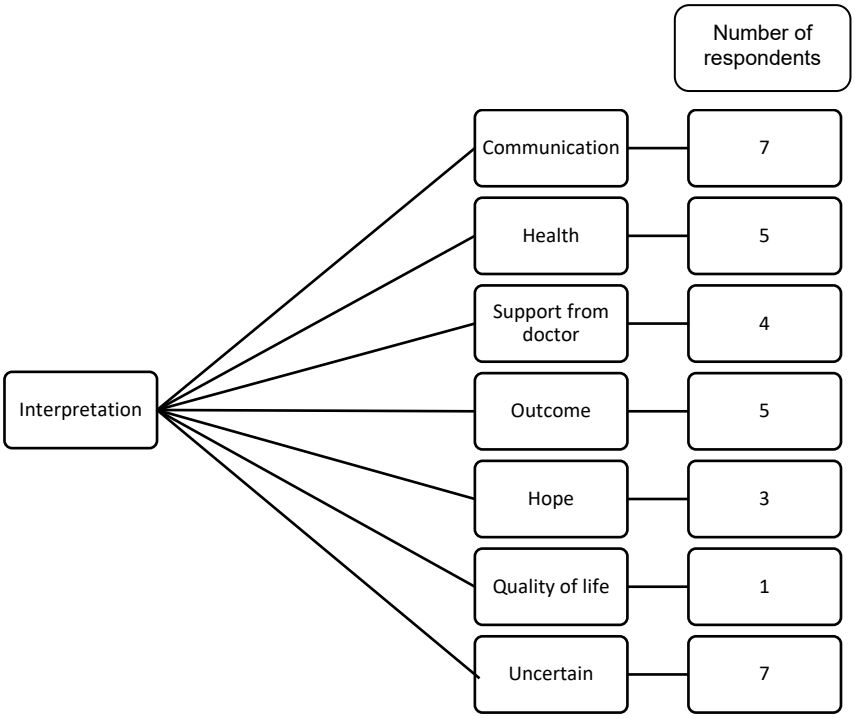
9.4. Other

Qualitative Interpretation.

Theme 1: Interpretation of the process domain.

Figure 25 summarises respondents' interpretation of the process domain. All respondents were able to complete the questionnaire and provide valuations for both the VAS and TTO health states. However, respondents' interpretation of the domain differed. Their interpretations ranged from the domain being an indicator of communication, to a reflection of the patient's quality of life, or outcome from their treatment.

Figure 25. Themes on the interpretation of the process domain.



As might be expected most respondents described the additional domain as reflecting communication, but from one of 3 different perspectives: the doctor and patient; the doctor's communication alone; or the patient's communication alone. The following quotes show some of the comments respondents made:

> "I feel like I wasn't good enough at saying telling what I feel like, is that right? And that the doctor didn't give the proper feedback maybe? (F6)

> "So you might come away thinking I wish I'd said that, or I wish I'd asked that, or whatever," (F5)

> "It could potentially be a doctor, very, you know well communicated things properly" (M7)

> "You feel like the doctor has listened to you, and understands what you are going through, um." (M10)

Others interpreted the additional information as an indicator of the support provided by the doctor. This is illustrated by the quotations below:

> "This person looks like they are not getting the help they need" (M1)

> "The doctor is not on your side" (F2)

Five respondents saw the process domain as a reflection of the patient's health, or the likely outcome from their illness. The quotations below illustrate this:

> "The doctor say like communication experience is poor, so he suggests that you know, I have problems." (M8)

> "If I go to the doctors for a health issue even just seeing the doctor if I've had a good experience, I feel like it's on the way to being sorted." (F1)

> "You know there is a chance that you can get better." (M2)

Although all respondents completed the questionnaire booklet, some were unsure of how to interpret the process domain. This is indicated by the questions they posed below:

> "Can I just ask when it says communication experience, what does that really, what does that mean?" (M2)

> "Can I just ask about the um, consultation thing, um, when, when its evaluated as communication experience as good, does that mean good news was given, or just kind of like you enjoyed having a chat." (M3)

"I was a bit puzzled by that really……I wasn't too sure what conclusions to draw from that really." (M7)

Theme 2: Impact of having a consultation.

The second theme which emerged from the data related to the impact the positive or negative communication experience had on the patient. Just as the interpretation of the domain differed between respondents, it was clear that they had different opinions on the impact the process domain would have on the patient. Although some respondents felt that a negative communication experience would have no impact on the patient, others felt it could impact on the patient's mood, the way that they feel, their ability to cope, or their health. Seven of the 20 respondents verbalised that a negative consultation would impact on the way the patient felt. Feelings included despair, anger, downheartedness and disappointment. The following quotations illustrate this:

"You'd feel more rubbish, um, if you didn't feel the doctor was supportive" (M10)

"The consultation with the doctor, it didn't affect the quality of life, although it probably would have made you feel a bit low." (F8)

"The communication experience was poor, I think that would make me feel worse about the situation, even though even though, you know, my health looks the same as it was on the last question. But I think it's just, ye, that that would feel um, you might feel quite upset about it." (F3)

Respondents also felt it may impact on a patient's hope, and one respondent felt it would impact on the patient's health.

> "Because then that hope's gone then for someone." (F2)

> "Not having had a good communication good with your doctor whether it's how you have portrayed it or he has received it sort of doesn't um, give you any hope or anything like that" (M2)

> "I've come out and it had impacted on the way I'd felt and perhaps hindered my um, kind of improvement in health" (M10)

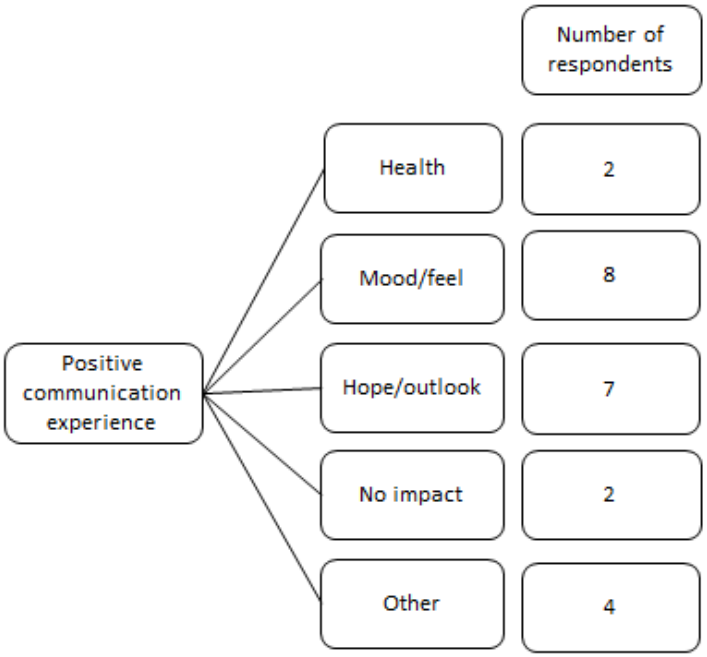
Several respondents felt that if they had a poor communication experience with their doctor, then they would seek a second opinion, and therefore it would have minimal impact on their valuation of the health state:

> "Perhaps I would try another doctor if I wasn't happy or something like that I mean I wouldn't think it was the end of the world, that's for sure." (M3)

> "Well my attitude, toward doctors like um, if I'm not satisfied with one, I would again try another one" (M4)

> "In that case you might just try to see another doctor to get a second opinion." (M7)

Figure 26. Themes on the impact of a positive consultation.



Respondents also felt that a positive consultation would impact on the patient; the sub-themes identified can be seen in Figure 26. Again this included an impact on the patient's mood and their hope and the outlook for their illness. The quotations below provide some examples of these:

"It gives you hope" (M6)

"You know there is a chance that you can get better." (M2)

"Being able to have good communication both ways with the doctor, um, does have an um, a profound effect on how you feel about um, your health state." (F5)

"If I go to the doctors for a health issue even just seeing the doctor if I've had a good experience, I feel like it's on the way to being sorted." (F1)

Two respondents felt that a positive communication experience with the doctor would have no impact on the patient:

> "If you have a good consultation with the doctor it's not going to do anything to change your health." (F8)

> "I suppose you feel better than when you don't feel like um the doctor's understood what's going on. Um, but then, I don't think overall, it would have that much impact." (M10)

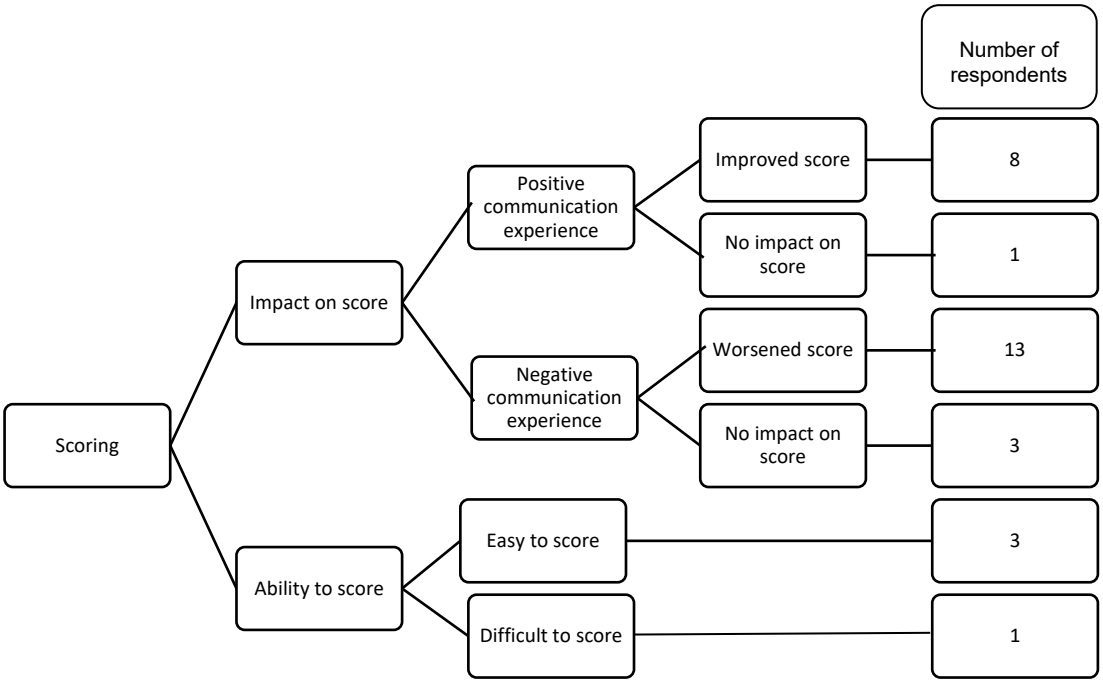
One respondent felt that the impact of a consultation would be dependent on the severity of the underlying health state, with a negative scenario, having a negative influence on an already poor health state:

> "But I guess, it is more significant if you are feeling tense already and then you go to the doctor and don't get anything from that, and it is a very poor consultation. Then it adds there. But if you are feeling okish and you have a poor consultation. I still think you would get over it." (M1)

<u>Theme 3: Scoring of health states.</u>

When valuing health states using the VAS, they must be scored on a scale anchored by 0 (dead) and 1 (full health). The aim of this study was to understand how respondents scored health states which included a process domain. For example, what impact the addition of the domain had on their valuation of the health state. Figure 27 shows the impact the addition of domains of different severities had on respondents scoring of the VAS.

Figure 27. Themes on the impact of a positive consultation.



All respondents were able to value the with-process health states using the VAS. For those who felt the process domain would affect their valuation, the addition of a good process domain improved the score:

> "This question appears to be the same as the other question apart from that the consultation with the doctor was good. So I'll put this back up to 80 because it would make you feel more positive about your outcomes and how things would progress." (F8)

> "They might have been referred on for some sort of treatment towards what's wrong with them so I would rate that more highly, 85 probably, because you tend to have a more positive outlook on your health state if you've seen a doctor and have a good experience than if not" (F1)

> "Um, I would say 50 in the last one, because you know the doctor say um, and communication experience is good. So it may not be as bad as I think." (M8)

One respondent reported feeling that a negative consultation would have no impact on the value of the VAS health states:

> "So I think the only difference is the um, consultation with the doctor. So, for me it's very similar," (M3)

And no respondents felt that the addition of a good communication domain lead to a reduction in VAS scores.

When valuing a health state which included a poor communication domain, the most common response, made by 13 of the 20 responders, was that the score reduced. Thirteen of the 20 respondents felt that the addition of a poor health state would result in a lower valuation than its paired SF-6D state.

> "I suppose it's the, it's the same as the last one except you've just had a rubbish experience with your doctor, and how that's made me feel, or you feel, gloomier about your, the outcome or, gloomier about how you might deal with the problems you're facing so, um. Ye, that might not be a long term kind of impact on the health state, but it can have an um, impact, on initial moods, um, so I'll put that a little bit lower." (M10)

> "So I'd value this a little bit lower, because that would make you feel a bit downhearted about this, and less positive about things" (F8)

> "um they've had a poor communication experience with the doctor. Um. So I'd value this a little bit lower, because that would make you feel a bit downhearted about this, and less positive about things." (F8)

> "Where it says I had a consultation with the doctor and the communication experience was poor, I think that would make me feel worse about the situation, even though even though, you know, my health looks the same as it was on the last question." (F3)

For 3 respondents, the addition of a poor process domain had no impact on the score:

> "I don't care about my communication with the doctor. Um, it limits social activities, so again, I think I will give it a 70." (M5)

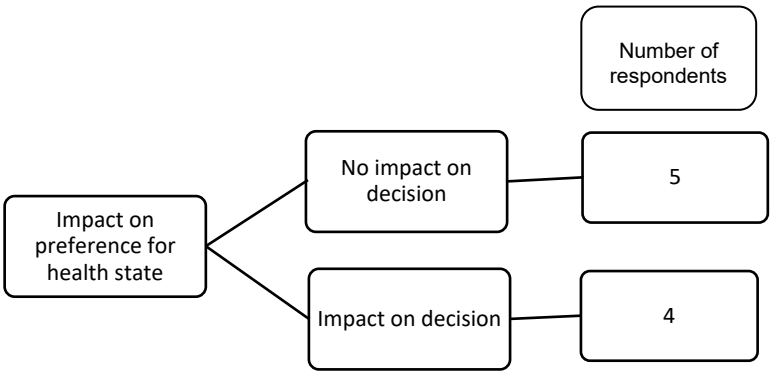
Four interviewees indicated that the process domain impacted on their ability to score the health states. For three, (F9, M1, M2), the addition of the domain process domain made this task of valuing the health state easier, and for 1 more difficult (M8).

Theme 4: Impact on preference for health states

When using the TTO technique to derive utility values for temporary health states, respondents are initially presented with a choice between 2 alternatives of certainty. They are asked to choose between a fixed duration of time (*t*) in imperfect health (the health state of interest), or a fixed duration of time (*x*) (where $x<t$) in a more severe health state (the anchor state), both followed by full health. They are then asked to choose between the anchor health state for time *t*, followed by immediate death, or a fixed time in full health (*x*) (where $t>x$). In both steps, the time is then altered until the respondent becomes indifferent between the 2 health states. The think aloud interview therefore aimed to evaluate whether the addition of the process domain to a health state impacted on a respondent's preference for

the health states when presented with these choices. Figure 28 shows the impact that the addition of the process domain had on respondents' preference for the health states.

Figure 28. Themes on the impact on preference for health.



Some respondents verbalised that the information provided within the process domain would influence their preferences:

> "Again, um, I feel that quality of life would be better um, for 2 years, then health state A for 5 years, especially, there's um, you'd feel more rubbish, um, if you didn't feel the doctor was supportive, um, so health state um, A for 5 years versus 1.5 years, so, I think full health for 1.5 would be better again because of quality of life, um." (M10)

> "Ye, it's the same principle I guess, and um, this time, this time the doctor was in it and it's a was the communication was good, and B was it wasn't it was very poor, so that has an added effect aswell." (M9)

There were also some respondents who felt that the process domain had no impact on their decisions:

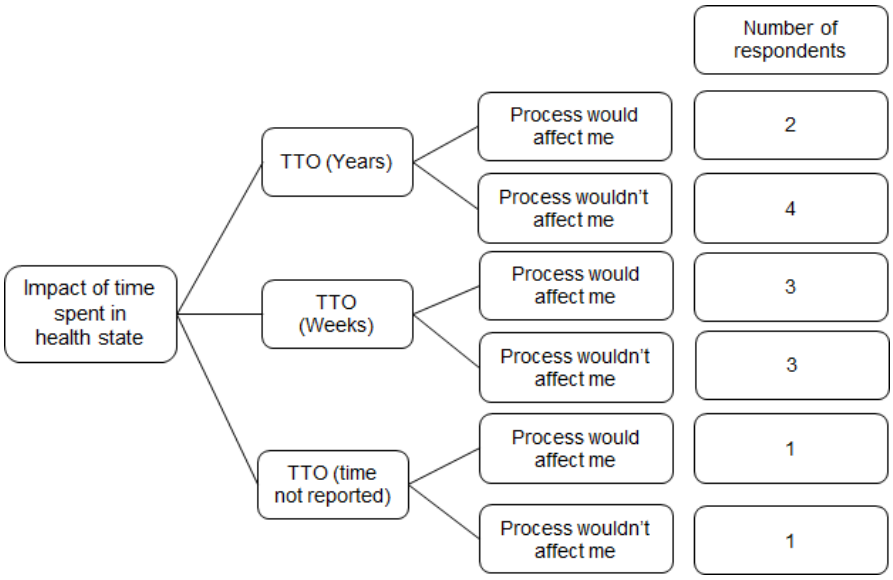
"no I don't see, I don't, I don't really, I wouldn't really get influenced by the type of communication I have by my doctor, and I don't think that, no I'd still feel the same things for 4 weeks, 1 week." (A4)

"Because it won't affect me, it wouldn't affect me how I talk with the doctor." (M11)

Theme 5: Valuation Method

Respondent's views and influences on valuing the combined health and process states differed according to which valuation method was being used. One difference, identified by respondents, and a specific area of interest for the researchers, was the time spent within the health states. This study hoped to ascertain whether the addition of the process domain fitted into health states being valued, and the concept of "time", particularly given the duration of a consultation and of the effects felt by the patient. An overview of the key related themes can be seen in Figure 29.

Figure 29. Themes on process and time.



When choosing between health states during stage 2 of the TTO valuations, some felt that the "concept of time" meant that process would not impact on them:

> "I guess, I'd maybe putting it into the context of time, thinking well could I cope with a crap doctor for 5 years? But then, if you think about it, actually you cope, you'd probably change doctor if you felt that strongly about it, so maybe contextualise it in terms of time." (F1)

> "the doctor maybe really negative, but then I'm also thinking like well, this is something about 5 years, or um, there will be many people around and, um, so, I wouldn't change much like, maybe (fills in the table), ye, the doctor wouldn't change it." (M4)

Two respondents discussed this specifically relating to their time alive, indicating that these respondents were not willing to trade time alive for a better communication experience with their doctor.

> "So, if the doctor treat me very poorly, would it affect my, I mean my decision to live? Um, no I wouldn't let him. I would still do what I want to do. (M11)"

> "It's hard to say what exactly happens and stuff, but if I'm in a state where I've only got 5 years to go, or I've got a quality of life for lower years I don't think the doctors kind of opinion would make a difference. " (M9)"

However, others choosing did feel that process would still affect them:

> "I think it was definitely more important if it's followed by death. I think that's really important that you are having a good communication with your doctor. Um, in the other, um the other one was where you compare 2 health states and which one you would rather be in. ye, um (silence for a few seconds, thinking) and it was only ever a month so it wasn't, again I don't think it was significantly significant for a month. Whereas the other one where you are talking about much greater expanse of time, and is followed by a definite death, I think it's really important that you have some encouragement, or that you feel like you are going somewhere with your doctor. Coz if not, that just, ye, can tip you over the edge a lot, so ye." (M1)

> "Because. Ye, is the first one where you are going to recover regardless. So I think. I know 4 weeks might sound not very much for someone who hasn't personally experienced such bad health problems. It's probably going to feel like 4 years. But still its only 4 weeks and then you'll be back to normal. Whereas this one, it's very extreme, so ye, I think that in question 2, the doctor consultation would make a difference. Whereas for the first one, no." (M6)

When considering the temporary TTO, for durations from 1 to 4 weeks, there were the same contrasting views. Some felt that the duration of the health states, would not impact on their preference:

> "It wouldn't really affect me how the doctor talked to me in a sense because like you are having a longer period of time, I mean if you are going to see the doctor just for one, just for one occasion then probably he might offend you because you can't like keep in your mind, but in this case you are living in the disease for like 4 weeks, which so I suppose it wouldn't matter what the doctor says, if it were me you know. OK, so in this case." (M11)

And others felt it would:

> "I think when maybe 2 weeks or 3 weeks and you've had no um, sort of positive consultation with your doctor, um, then its going to start to, um, make it a lot more different to a, um, make it a lot more serious than A" (M2)

For the VAS valuations, no timeframe was reported. Some, respondents still considered time when scoring the health states, and made assumptions to aid in their valuations. This is discussed further in Theme 8.

> "I think so, but in that case, because of the one occasion, it would affect me, I mean personally a bit more than this." (M11)

Figure 30. Themes on the ease of completion according to valuation method.

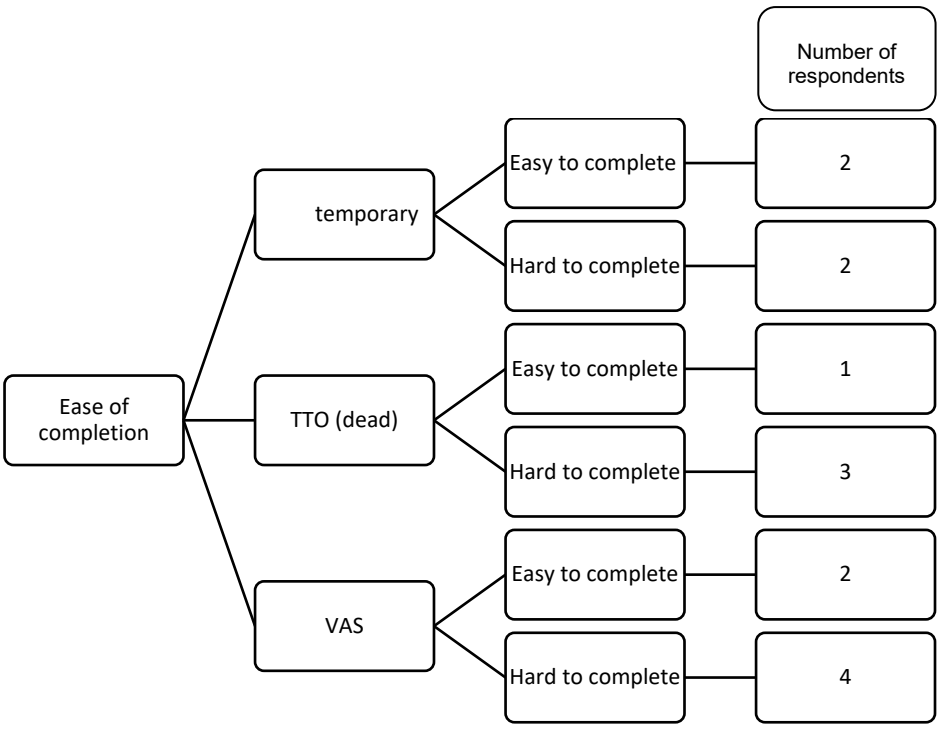


Figure 30 provides an overview of the respondents' ease of completion for the different valuation tasks. Despite the increased complexity of the TTO questions, when considering the health state with process, respondents overall found these health states easier to value than the VAS:

> "I'm not quite sure why whether it's a visual thing I'm not sure. Um. I don't know, um these were just, these were just easier to compare, um. I don't know. It's just easier to compare and read through than, um, those." (F9)

Although some respondents also stated that both types of TTO questions were most challenging:

> "um, ones at the back, um, just trying to figure out where I stood I suppose against death and time, and what's better, so I suppose just trying to

formulate where I stood on it, that took a little bit of time, when you're faced with kind, of, kind of the death and full health and stuff, so." (M10)

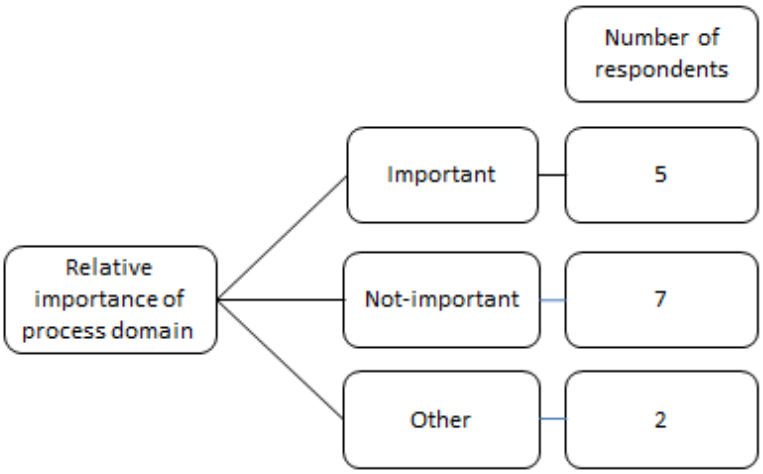
"Um, I think I struggled more with the time alive than the other ones. The other ones felt quite, they felt easier to mark on the scale. I found that one harder, um, but it was really interesting to think about kind of in terms of um health state and then death." (F3)

"I'd say the hardest one was towards the beginning of the second booklet. Yep, this one, ye, this one was mind blowing a bit at times, the first question…… It's because of these middle ones because of the 2 and 3 weeks and 2 and 4 weeks, and A and B both having their pros and cons if you could say. So ye, um, ye, so as you can see, I have ticked and crossed the wrong boxes by accident." (M6)

Theme 6: Importance of process domain

During the semi-structured interview, respondents were asked, when looking at the health states, which aspects were most important to them when coming to their decisions. A summary of the responses can be seen in Figure 31.

Figure 31. Themes on the relative importance of process domain



One respondent felt that the consultation with the doctor was the most important.

> "I think it is very, very important, probably more so than the rest of the things." (M2)

Others felt that, although not the most important domain, it did remain important to consider:

> "I would rate it fairly high." (M6)

Although some respondents felt it was the least important domain:

> "This one for sure. But I always put like, symptoms are the way I feel. Like they are always a priority, I don't think this doctor thing has anything to do with these answers I am giving." (F6)

> "Doctor wouldn't be the main thing for me, it would be the other things that I'd prioritised."(F1)

"And ye, I'd say the consultation with the doctor is last." (M9)

Despite the above responses when specifically questioned about the importance of the domain, throughout the think aloud interviews more respondents stated its importance:

"I mean I think that it is major, because I know myself, if do go to the doctor, and I feel I've been able to express what I want to say well, and um, I have a fantastic GP, and I just always feel better (giggles) when I've seen him." (F5)

"Again, I think um, the stuff about having a consultation with a doctor, the expert, on what you are suffering, is the key," (M2)

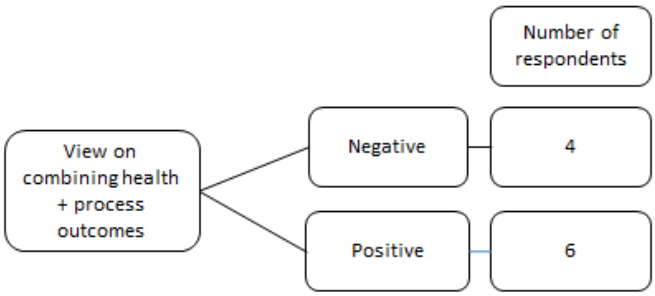
Some were not sure, or felt it was dependent on the valuation method or severity of health state being considered:

"Well in terms of comparing 2 health states I think it made a difference. But when you're comparing to full health, I don't think it did make such a difference." (M10)

"I mean I don't know based on its context how treatable their problem is? So the difference between having a doctor and not having a doctor might be minimal, or it could be a big difference." (M6)

Theme 7: Combining health and process outcomes.

Figure 32. Themes on combining health and process outcomes.



A summary of the key themes discussed in response to health scenarios which contain both health and process outcomes is provided in Figure 32. Respondents were able to read the SF-7D health states and value them using the 2 valuation techniques, and it was evident that some made a clear distinction between the health and process components:

> "The consultation with the doctor, it didn't affect the quality of life, although it probably would have made you feel a bit low." (F8)

> "I mean it's a plus if you have a good consultation but I don't think that it can't all be good and if you've got a poor health outcome anyway. If you have a good consultation with the doctor it's not going to do anything to change your health. You can't do anything about your health anyway." (F8)

However, when asked specifically what their views were on combining the 2, there were both positive and negative responses.

Some respondents had a negative response. The comments included:

> "But ye, doctors like individuals doctors are individuals and they might have had a bad day, and just made that experience poor for that patient. So, I guess that's maybe something to, account for." (F1)

> "The doctor? No, basically I care as much as if a sales cleric has been nice to me. Why would I be bothered?" (M5)

Others responded positively:

> "I think it gives people a different view, definitely. Because I think, if that's not there. Most people would just assume you would have a positive experience, or you could go and see someone, or something could be done, you know. But I think if that's put in place, and it's, you know." (F2)

No respondents reported any negative effects of including the process outcome, even those who placed no value on it at all.

Theme 8: Heuristics

Within this context, heuristics refers to mental shortcuts that ease the cognitive load of making a decision. One example arose in response to the VAS questions, where respondents made their own assumptions to fill in this data gap. In this case, the respondent made the assumption that it was a "one off" disease, which was only suffered on the "one occasion" (M11).

Another respondent used heuristics to develop a methodology for valuing the health states without having to process all of the domains and their severities. She used the process domain as a proxy reflecting the overall status of the patient given the health attributes listed previously. This is illustrated by the quotations below:

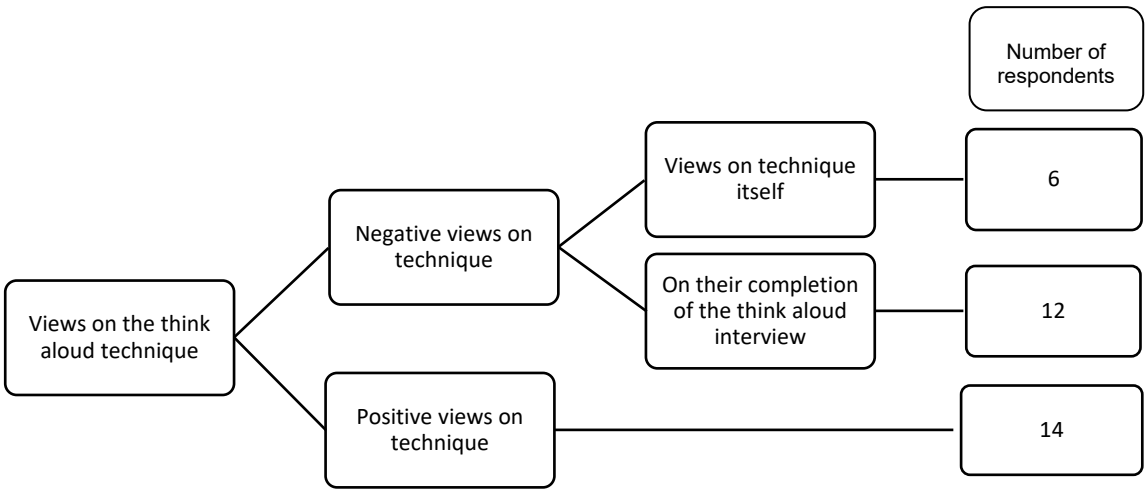
"I've jumped to the last line as well and it says your communication experience was very poor, so obviously it's not looking quite as good." (F9).

"I'm I think as I was going along I was sort of jumping to the top and then jumping to the bottom to see what the outcome was, a little bit more for the doctor. And thinking oh alright, this isn't going to be as well, or this is going to be better this one. So it sort of surmised it slightly." (F9)

Theme 9: The Think Aloud Technique.

As well as exploring the health state valuations themselves, this study also sought to determine whether the think aloud technique itself impacted on participant's responses, and valuations of the health states. At the end of the interview, all respondents were asked for their views on the technique. Both positive and negative feedback was received. An overview of the subthemes can be seen in Figure 33.

Figure 33. Views on the Think Aloud Technique



Almost three quarters of all respondents had positive comments on the think aloud technique:

> "I think it's ok. I really enjoy it" (M8)

> "It's good. If only I was allowed to do it in an exam. Ye, and I think ye, because when I speak out loud. I mean I cannot listen to it, if I think silently I mean really quiet I get distracted by things around me, like if there is a clock ticking." (M11)

> "I think possibly it may have, it may have helped me um, think about things, a bit more laterally and a bit more holistically really." (M7)

The negative feedback was both on the technique itself, but also reflected respondents views on their own completion of the think aloud interview:

> "Just personally I just find it normally when I think about things I am silent and I am thinking about things for a long time, so it is completely the opposite to how I normally think, so um, I found that kind of made me do something which I was quite uncomfortable with," (M2)

> "If I was to assess myself I'd say I didn't think out loud enough, I think. Well it was my struggle with it was just getting my head around the, the way that the question worked, and so I was, I don't know, kind of, I don't know going around in circles a little bit trying to work it out." (M3)

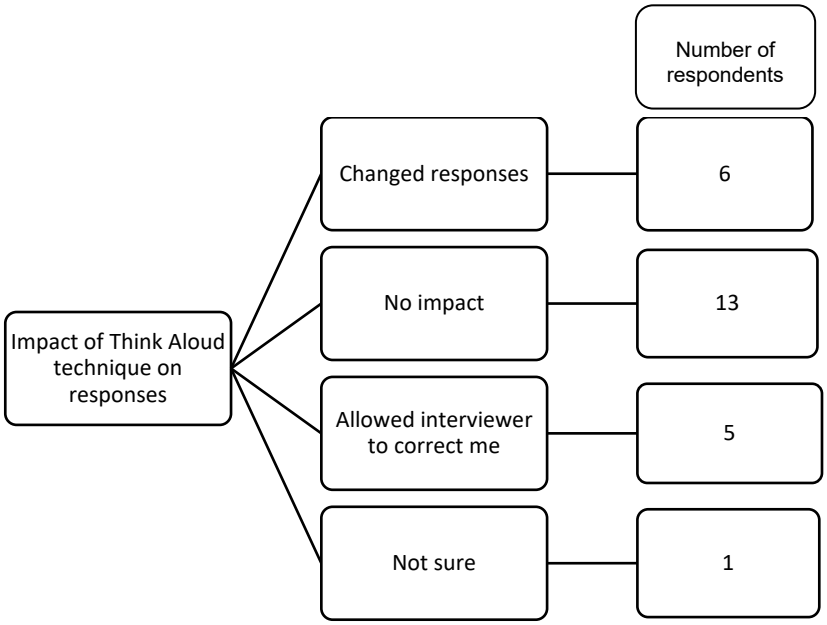
> "It's kind of strange at first. Kind of bombarding you with my thoughts, and should I keep them to myself?" (F7)

> "I think it does detract, um as it's just shown, I don't think you read the question properly. Um, it's like in an exam where you read the question properly, and I think it detracts when you're trying to think aloud." (F8)

> “Really weird. It doesn’t feel right. I don’t know. I mean. It feels like I still have thoughts inside my head that I’m not saying out loud. It feels as though my voice is clashing with the thoughts inside my head. And I mean t hats the best way to describe it.” (M6)

There was also a specific question posed to each respondent, “do you think, if you hadn’t been asked to think aloud your answers would have been the same or different?” The responses were varied. An overview of the subthemes is illustrated in Figure 34.

Figure 34. Impact of Think Aloud Technique on Responses.



Six respondents felt that thinking aloud did change their responses. Of these, 1 thought it had a negative impact. Their reasons for this are described in the

quotation below.

> "Um, maybe would have spent some more time thinking. Hm, maybe having to verbalise everything made me rush through it." (F4)

Two felt that thinking aloud lead to better constructed responses:

> "It was nice knowing that actually I am being asked to take my time and think and go through the questions. Especially as it is quite a hefty thing. It was, ye, I didn't feel pressured, it was like that was what was expected, and it was nice that I could take that time so ye. I think my answers are more thoughtful and more, ye accurate." (A5)

Despite these negative views from some, respondents said the TA procedure had no impact on their responses:

> "I mean the same would have happened if I'd been thinking in my head." (F7)

> "What I'd say is that everything I said would be the same, thinking out loud or not thinking out loud, on the assumption that the person I am considering is somebody who sees the doctor as the only source of information or the best source of information." (M2)

A further question was answered to determine whether respondents felt uncomfortable with the interviewer being present and listening to their thoughts. Again, over half of respondents (n = 14) reported that the interviewer listening had no impact on their responses, although 4 did feel their responses would have been different if alone and thinking aloud:

"Well I think always having somebody in the room, you are all. I am very much aware of my surroundings all of the time. So even though I feel comfortable and relaxed. It couldn't help but influence it" (F4)

"So if I was in a room and just kind of talking? I think that would have been different" (F3)

7.3.3 Think Aloud Discussion

This qualitative study to explore respondents' reactions to health scenarios which combine both health and process outcomes has furthered the earlier quantitative work which involved their development and valuation by providing us with a more in depth insight into the decision making process of 20 respondents. The study used the think aloud technique which allowed a detailed investigation of how respondents completed the valuation tasks, the factors they considered, and their decision making processes. The study found that although participants were able to complete the valuation tasks, there were several findings linked to the way they completed the tasks, which may impact on the resulting valuations. This study has also added to the published literature in the field, being the first qualitative study to explore health states which contain both health and process attributes, and using and describing in depth, the framework approach for analysis.

All 20 respondents were able to work through the questionnaire booklet and provide the information required to calculate the valuations, providing an initial indication that it is possible to combine health and process outcomes, and ask the general public to value using the VAS or TTO. This finding is in line with the previous studies. It could therefore be assumed that the addition of a process

domain does not make the valuation task too complex. However, this study did highlight some significant concerns regarding the validity of the valuations.

The key concerns, which will significantly impact on the validity of the valuations, relate to respondents' understanding and interpretation of the process domain. The domain stated, for example: "You have just had a consultation with the doctor. You evaluated your communication experience as poor." Although some respondents understood this as being directly related to communication, there were others who made statements indicating they did not understand its meaning. The remaining respondents made alternative interpretations of the information, for example, they saw it as an indication of the level of support provided by the doctor, or as a reflection of the patient's health or potential outcomes of care. These issues have been identified in previous studies, For example, as cited by Al-Janabai et al. (2013), Tourangeau et al. (2000 p.117) suggest, in their work into the requirements for questionnaire completion, that "in appropriately answering a question, and individual must … understand (comprehend) the question in the way that the researcher intended;". Al-Janabai et al (2013) explored this "error" through the application of their coding framework in their Think Aloud study into the completion of a self-capabilities questionnaire and the EQ-5D. Their study found that, out of 34 respondents, there were 2 comprehension errors on completing the ICECAP-A measure, and 2 when completing the EQ-5D. Although our coding framework differed, there were 7 respondents who questioned the meaning of the domain, and others, who although they felt they understood the process information, made questionable interpretations, including it being a reflection of a

patient's health status or an indicator of prognosis. We could assume this leads to uncertainty in respondents' comprehension of the health states, and therefore questions their ability to value them credibly. This seems particularly important as some of the interpretations were so far from the intended meaning, and are likely to have a large impact on the valuations, for example, those who felt that the process domain reflected the patients' outcomes of care.

Despite being provided with the same information, respondents brought different information into their decision making process, for example, new meanings to the process domain, e.g. it is a reflection of hope or prognosis. They also placed different emphasis on the different attributes in the health and process states. This reflects findings by both Baker and Robinson (2004 p.38) and Mulhern et al. (2012). Baker and Robinson (2004) examined responses to SG questions, and asked "What information do respondents consider when asked preference elicitation questions?". They identified themes such as "considering others" and a "desire not to become a physical burden". Mulhern et al. (2012) explored valuations of TTO and DCE health states, and identified that although some respondents thoroughly considered all attributes, others focused only on selected aspects of the health states, for example, the first, or last. This was certainly evident within this study, and is illustrated by the respondent who used the final "process domain" as a proxy for the patients overall health status. Potential further work could therefore include health and process attributes, and the placing of the process domain within the list of attributes included in the health scenario's.

The use of a "short cut", as illustrated by the respondent above who used the final process domain as a representation of the patients overall condition is an example of a "heuristic" – a mental shortcut that eases the cognitive load of making a decision (Blumenthal-Barby & Krieger, 2015). As detailed in Theme 7, there were some, but only minimal examples identified in this study. In addition to the respondent who used the process domain as a proxy reflecting the overall health status of the patient, another respondent applied their own assumptions on time constraints to the health states where this information had not been provided. As the use of these strategies in answering health state preference studies is utilised by some respondents, and is likely to impact on the resulting valuations, this is an area which should be researched in more depth in relation to the preference based health state valuation methods. At present, the bulk of the research into the area is focused on DCE studies (Ryan et al. 2009), despite the dominant and preferred utility elicitation methods being TTO and SG.

Not only did the interpretation of the process domain differ between respondents, so too did the perceived impact that both a positive or negative experience would have on the patient. For example, several respondents commented that a positive or negative communication would affect the way the patient felt, and also the hope they had for the improvement of their condition. This finding is not wholly comparable to other studies, as none considered process attributes, but it is important in that it indicates that patients' experiences of different processes of receiving care can affect the way patients feel, and therefore should be considered in resource allocation and cost-effectiveness decisions. In addition to this it

suggests that respondents bring additional information to their decisions, above the information provided in the health scenario descriptions. This finding is in line with previous work, for example, Mulhern et al. (2012) found that respondents completing TTO and DCE valuations incorporated a range of personal factors linked to health and life experiences into their decision making process, and Baker & Robinson (2004) found that respondents completing the SG task considered non-health factors such as the impact of choice on others. This introduction of assumptions and information outside of the descriptions provided is certain to influence people's valuation of health states.

This study identified a further concept, which is not relevant to the usual health domains included in health state descriptions, and therefore was not identified or considered in previous work – seeking a second opinion. Seven of the 20 respondents felt, if presented with the situation where they had had a poor communication experience with their doctor, they would seek a second opinion. They therefore placed less emphasis on this domain than those specifically relating to their health, and in some instanced disregarded it completely. This has serious implications for the addition of such a domain, which differs from the health attributes in that alternatives can be sought out.

Overall, it appeared easier for respondents to value the health and process states when using the VAS. The findings from the study also indicated that the addition of the process domain can impact on respondents scoring of a health state using this approach. Given the complexity of the TTO exercise, it was harder to identify through the qualitative research how much influence the additional process domain

had on respondents' preference for a health state, than to determine the impact on its score using the VAS. However, some respondents made clear statements illustrating their views, and they did find it feasible to consider the communication experience in their decision making. They were willing to trade longevity for a better communication experience with their doctor. Again, this finding cannot be compared to previous work, as this is the first think aloud study which explores health scenarios including process attributes.

The study also provided additional insight into the impact of the duration of the health states. This is particularly relevant to this study, as whereas a patient's condition may be reflected by the health domains within the health state for a long duration, the consultation with the doctor only represents a small piece of this time. Therefore, given the short duration of a consultation, and considering how long the impact of a consultation may last for, we hoped to ascertain whether a 4 week duration for the temporary health states, and 5 year timeframe for chronic states followed by death were realistic and sensible assumptions to make. The study showed that although some felt the time parameters meant the process domain would not influence their preferences, for some it was still seen as an important component of the health and process state to consider. Several of the studies identified for the review considered issues surrounding the duration of health states. For example, Spencer (2000) only included patients up to the age of 59 years, to ensure that all could imagine still being alive after the 10 year health state duration – alternatively they could have reduced the duration to 5 years.

However, the findings from this study indicate that the durations included in this study were plausible assumptions to make.

Limitations and strengths

This study does have its limitations. Although the sample was loosely based on age and gender similar to the general population, it comprised of predominantly university workers and students, with only 1 of the 20 respondents being retired. Despite attempting to select respondents who reflected the general population, several potential participants were unable to attend interview dates and times, or did not turn up to appointments. Therefore, the ability to select purposively based on age and gender was limited in practise, and significantly influenced by the number of available respondents. The sample was also highly educated with 10 being educated to degree level. There was also a high proportion of participants who had never been married (n = 15), and the mean age was 30.8 years, indicating that it is likely the sample were also relatively healthy. Therefore, despite attempting to select a population which was representative of the general population and which we may have more confidence that the findings are generalisable, when reviewing the results these population differences should be remembered and their impacts considered. For example, older people may value health states using the TTO approach differently from younger people, given that the former have fewer years left to live. The imminence of death may distort the valuations. (Cook et al. 1994). More healthy participants may also value health states differently to those who are unwell, as it is likely that their experiences of illness will impact on how they value health. If further work were to be undertaken,

it would therefore be appropriate to ensure a sample more representative of the general population.

It could also be suggested that the inclusion of the TTO in the think aloud study could have taken away from the findings from the VAS, particularly given the original incentive for the study was to understand more fully the findings from the process utility valuation study reported in Chapters 5 and 6. However, I would suggest that its inclusion enhances the findings from the study. It is accept that the study stemmed from the high number of irrational responses to the process utility valuation study. However, if utilised more widely, the health state valuations would need to be based on a preference based measure such as TTO. Hence the decision was made to include TTO in the think aloud study to allow us to identify and understand the implications of combining health and process outcomes using a preferred valuation method, with the aim of producing findings which were more valuable than including only the VAS. It provided important information on whether respondents were willing to trade time with ill-health and time alive for changes in severities of process utility. This information would not have been obtained if only the VAS was included.

The potential that the impact of the process domain is also related to the severity of the underlying health state was not explored fully within this study. Only one person suggested that the impact of the consultation would be dependent on the severity of the underlying health state. This is in alignment with the findings from Brazier et al., (2011) in their study assessing the impact of adding a pain dimension to a preference based measure, although was not supported from the process

utility valuation study, where the regression model incorporating both the process domain level and severity of the underlying health state was non-significant. If taken forward, further work into adding a process domain to existing preference based measures should consider this issue further, and understand more fully the impact of its inclusion on the co-efficients of the health domains. Due to the concern not to have too high a completion burden on respondents, this study only included clearly positive or negative communication domains. The inclusion of the "average" communication domain was not explored. In order to fully understand the impact of the process domain, this should be explored further.

There is also a concern that the findings from this think aloud study may contain focussing effects. These may particularly be present as respondents were provided with pre-study information, explaining the concept of health and "process" outcomes, and that the aim of this study was to explore their reaction to adding a "process domain". It is likely that informing respondents of this aim will have impacted on the results, potentially drawing their focus more directly to the process domain, and emphasising their reactions. This is in addition to the obvious difference in health states between those with and without the process attribute. In hind-sight, it would have been more appropriate not to have informed respondents of this specific aim, as this would have identified more realistic reactions and thoughts to the process domain.

Therefore, as well as further quantitative work, there may be an indication for further qualitative research into these combined health scenarios, which is specifically designed to remove this focusing effect, for example, including only

health and process states. Although this concept was identified as an issue in the bolt-on studies reported in Chapter 5 (Brazier et al., 2010) it was not seen as relevant to the think aloud studies which explored only standard preference based measures.

There are also the issues related to the technique itself. Although the researcher sat slightly behind the respondents, to ensure they were out of view, respondents did ask questions when faced with a question they did not understand. This follows the approach taken by Cheraghi-Sohi et al. (2007) where the researcher only provided prompts and responded to occasional requests for clarification. Within this think aloud study I refrained from imposing my view on the respondent, but responded by saying that one of the aims of the study is to understand how you interpret this aspect of the health state. So if you imagine you have just come out from a consultation with the doctor and you evaluated your communication experience as "poor". What would you interpret this as meaning? However, the very presence of a researcher is likely to influence participants' responses to some degree. And any involvement or responses to questions may introduce bias. This introduces the concept of reflexivity, whereby the researchers past experiences, assumptions and preconceptions can affect their qualitative research (Bryman, 2012). The use of concurrent think aloud techniques is purported to reduce the levels of bias introduced by the researcher as they are primarily a "silent" observer of the task (Miller & Brewer, 2003), however it is likely that my past experiences of research, and associations with health care which includes 10 years working as a

physiotherapist will have influenced my work. This is inherent with qualitative research but should be acknowledged.

Despite the limitations, this study also has its strengths. It is the only think aloud study performed which asks respondents to value states which include both health and process outcomes. This expands on the current literature and provides additional insight into the concepts people consider when completing both the health and process domains within health state valuation studies, and also identifies some additional considerations for process, for example, the ability to obtain a second opinion and impact this may have on valuations.

This study included additional questions on the technique itself. Although other studies recognise potential limitations of the think aloud technique, for example, Mulhern et al. (2012) discuss that we cannot be sure if respondents are fully verbalising all of their thoughts, this study asked specific questions in an attempt to gauge the degree to which this may have impacted on responses. The additional question asking whether their responses would have been different if the researcher had not been sitting in the room explored this further. And for some, it was clear that the mere presence of another person did affect their responses. If this approach is to be used further this should be investigated in greater depth, and methods for addressing this limitation should be sought.

This study also applied the framework approach, and was explicit in reporting the steps undertaken in the analysis. This transparency in the methodology improves

on the reporting of the existing studies, and makes comparisons of the findings with other studies easier and more credible.

7.3.4 Think Aloud Conclusion

This study has raised serious concerns regarding respondents' understanding and interpretation of the process domain, and this leads to a questioning of the validity of the health state valuations which have been derived in this study. However, despite this, it should not be forgotten that some respondents felt that the consultation with the doctor was the most important domain presented, and therefore although this may not be the optimum format or information to use to represent process, the consideration of patient preferences for both health outcomes and different processes of care should both be considered. The traditional and most used approach for the measurement of the benefits of health care is based on the QALY. At the moment this considers only health outcomes. By excluding utility gained from the process of receiving care, the utility estimates used may not truly reflect patients' preferences for treatments. This is particularly important for studies which compare interventions where the primary outcome is associated with process, as in the ePAQ EEACT, as opposed to those associated with health.

7.4 References

- Al-Janab, I. Keely, T. Mitchell, P. Coast, J. Can capabilities be self-reported? A think aloud study. *Social Science and Medicine* 87,116–122.

- Baker, R. Robinson, A. (2004) Responses to standard gambles: are preferences 'well constructed'? *Health Economics*, 13(1), 37-48.

- Ball, CT. Langholtz, HJ. Auble, J. Sopchak, B. (1998) Resource Allocation Strategies: A Verbal Protocol Analysis. *Organizational Behavior and Human Decision Processes*, 76(1), 70–88.

- Blumenthal-Barby & Krieger (2015) Cognitive biases and heuristics in medical decision making: a critical review using a systematic search strategy. *Med Decis Making*, 35(4), 539-557

- Brazier, J. Deverill, M. Green, C. Harper, R. Booth, A. (1999) A review of the use of health status measures in economic evaluation. *Health Technology Assessment*, 3(9),1-164.

- Bryman, A. (2008) Social research methods. 4th Ed. Oxford, UK.

- Cheraghi-Sohi, S. Bower, P. Mead, N. McDonald, R. Whalley, D and Roland M. (2007) Making sense of patient priorities: applying discrete choice methods in primary care using "think aloud" technique. *Family Practice* 24,276-282.

- Cook, J. Richardson, J. Street, A. (1994) A cost-utility analysis of treatment options for gallstone disease: methodological issues and results. *Health Economics,* 3,157-68.

- Damschroder, LJ. Roberts, TR. Goldstein, CC. Miklosovic, ME. Ubel, PA. (2005) Trading people versus trading time: what is the difference? *Population Health Metrics*, 3:10.

- Devlin, NJ. Hansen, P. Selai, C. (2004) Understanding health state valuations: a qualitative analysis of respondents' comments. *Qual Life Res*, 7, 1265-77.

- Dolan, P. (1997) Modelling valuation for EuroQol health states. *Med Care,* 35, 351–363.

- Drummond, MF. Schulpher, MJ. Torrance, GW. O'Brien, B. Stoddart, GJ. (2005) Methods for the Economic Evaluation of Health Care Programmes. 3rd Edition, Oxford, Oxford Press.

- Ericcson, KA. Simon, HA (1993) Protocol Analysis: Verbal Reports as data. Revised edition. Bradford Books.

- Horwood, J. Sutton, E. Coast, J. (2014) Evaluatingthe face validity of the ICECAP-O Capabilities Measure: A "Think Aloud" Study with hip and Knee Arthroplasty Patients. *Applied Research Quality Life*, 9:667-682.

- Kuusela, H. Paul, P. (2000) A Comparison of concurrent and retrospective verbal protocol analysis. The American Journal of Psychology, 113(3), 387-404.

- Lamers, LM. McDonnell, J. Stalmeier, PF. Krabbe, PF. Busschbach, JJ. (2006) The Dutch tariff: results and arguments for an effective design for national EQ5D valuation studies. *Health Econ*,15(10),1121-32.

- McGavock, ZC and Traeharne, GJ. (2011) Young Adults' Beliefs about People Living with HIV/AIDS and Rheumatoid Arthritis: Thematic Analysis of a Think-

Aloud Questionnaire Investigation. *New Zealand Journal of Psychology* ,40(3),71-78.

- Miller, RL. Brewer, JD. (2003) The A-Z or Social Research. Available at: http://srmo.sagepub.com/view/the-a-z-of-social-research/n117.xml Accessed on 18th October 2015.

- Mulhern, B. Tsuchiya, A. Brazier, J. Rowen, D. (2012) How do respondents perceive health state valuation exercises? A think aloud study investigating time trade off and discrete choice experiments. Paper for presentation at EuroQol Plenary Rotterdam.

- NICE (2013) National Institute for Health and Care Excellence. Process and Methods Guides: Guides to the methods of technology appraisal 2013. Available from: http://publications.nice.org.uk/pmg9 [Accessed 28th August 2015).

- Papageorgiou, K. Vermeulen, KM. Leijten, FRM. Buskens, E. Ranchor, AV. Schroevers, MJ. (2014) Valuation of depression co-occurring with a somatic condition: feasibility of the time trade-off task. Health Expectations. doi: 10.1111/hex.12303.

- Pope, C. Ziebland, S. Mays, N. Qualitative Research in Health Care: Analysing Qualitative Data. *BMJ* 320(7227)114-116.

- Ritchie, J. Lewis, J. (2003) Qualitative research practice: a guide for social science students and researchers. London: Sage.

- Robinson, A. Dolan, P. Williams, A. (1997) Valuing health status using VAS and TTO: What lies behind the numbers? Soc Sci Med, 45(8), 1289-1297.

- Ryan, M. Watson, V. Entwistle, V. (2009) Rationalising the 'irrational': a think aloud study of discrete choice experiment responses. Health Econ, 18(3),321-36.

- Shaw, JW. Johnson, JA. Coons, SJ (2005) US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care*, 43(3),203-20.

- Spencer, A. (2000) Testing the additive independence assumption in the QALY model. University of London. Working Paper, No 7.

- Spencer, A. (2001) The time trade-off method. An exploratory study. University of London. Working Paper, No 437.

- Torrance, GW. (1986) Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5,1-30.

- Tourangeau , R. Rips, L. Rasinski, K (2000) The psychology of survey response. New York: Cambridge University Press. Cited in: Al-Janab, I. Keely, T. Mitchell, P. Coast, J. Can capabilities be self-reported? A think aloud study. *Social Science and Medicine* 87,116–122.

- Van Osch, SMC. Stigglebout, AM (2008) The construction of standard gamble utilities. Health Economics, 17:31-40.

- Whitty, JA. Walker, R. Golenko, X. Ratcliffe, J. (2014) A Think Aloud Study Comparing the Validity and Acceptability of Discrete Choice and Best Worst Scaling Method. . PLoS ONE 9(4): e90635.

Chapter 8: Discussion

8.1 Introduction

This thesis focuses on process utility. After exploring the limitations of EEACTs and applying these to the findings from the ePAQ trial an additional limitation is identified. Whilst the ePAQ trial was generally well suited to an EEACT in terms of comparator, length of follow-up and lack of protocol driven costs, the standard economic framework for undertaking cost-utility analyses, which involves the calculation of a cost per QALY, is suited only to those studies which have an impact on health outcomes. The most commonly used and favoured preference based measures which are used to derive the utilities for input into the QALY calculation e.g. the SF-6D and EQ-5D, focus purely on health effects, and fail to consider other non-health attributes such as those associated with the processes of receiving care.

The ePAQ trial was therefore used as a vehicle for exploring process utility further, with the ultimate aim to use the data collected from within the trial to improve the relevance of the trial outcomes. The aim was to adapt the SF-6D to capture both health and process outcomes in a form that the findings could be incorporated into the QALY. This has resulted in a systematic literature review of the process utility literature, the development of a health state valuation study which bolts an additional domain onto the SF-6D, results of the ePAQ EEACT both with and without the consideration of process utility, and a qualitative study into patients perceptions of combines health and process scenarios.

Within this chapter I set out the main contributions made by the thesis, outline the strengths and weaknesses of the research presented and discuss recommendations for future research.

8.2 Summary of Contributions Made to Research.

8.2.3 Literature review into the limitations of EEACTs

A literature review into the limitations of EEACTs was completed. As the review was opinion based as opposed to being data driven, a citation search was undertaken. The searches were initiated from 3 key papers: Drummond & Stoddart (1984); Drummond & Davies (1991); O'Brien et al. (1996). The searches identified 39 studies which met the inclusion criteria. Publication dates ranged from 1984 to 2015, indicating the longevity of the debate. Most studies discussed one or all of the 7 threats to validity introduced by O'Brien in 1996 (Choice of comparison therapy; Gold standard measurement of outcomes; intermediate versus final health outcomes; inadequate follow-up and sample size; protocol driven costs and outcomes; geographic transferability of evidence; selected patient and provider populations). A further threat, linked to the source of funding, was also discussed by Friedberg et al (1999). Although this literature review identified several relevant papers, none reported a literature review which provides an update of the current status of the debate.

8.2.4 The ePAQ EEACT

Chapter 3 reports an EEACT performed in pelvic floor medicine. The ePAQ EEACT applies standard economic methods to a trial comparing the use of ePAQ online in

advance of a telephone consultation (intervention) to ePAQ online in advance of a face-to-face consultation (standard care). The cost-effectiveness analysis indicated that the intervention is unlikely to be cost-effectiveness at a willingness to pay threshold of £20,000. However, the study considered both the costs of the consultations, and direct and indirect costs incurred over a 6 month follow-up period. If considering the consultation costs only, the cost of the intervention was less than half that of standard care. As there was wide variability in the follow-up costs, which introduced a lot of uncertainty, it may be most sensible to ignore these wider costs and concentrate solely on the costs of the consultation.

As well as reporting the cost-effectiveness results, the limitations of EEACTs identified in the literature review reported in Chapter 2 were applied to the ePAQ trial. This identified a further limitation. In cost-effectiveness studies which compare interventions where we expect no differences in health outcomes between groups, but we do expect differences in patient satisfaction, the standard economic methods applied will not capture all of these relevant and process orientated outcomes, as the QALY, in its current format, is designed to capture only health differences. This was clearly evident in the ePAQ EEACT, where the quality of life of patients as measured by the SF-6D reduced slightly for the intervention group, despite there being significant differences aspects of patient satisfaction as measured by the PEQ.

Although the ePAQ trial applied only standard economic methods, it did identify a novel perspective on process utility, and raised the question – should the utility function be expanded to include not only individuals' preferences for health

outcomes, but also their preferences for the processes associated with receiving care. Although this thesis does not attempt to answer this question – it does highlight that others have argued for the inclusion of process in the utility function, and therefore the incorporation of process utility within the QALY is legitimate. The issue of what is a valid utility function is beyond the scope of this thesis.

The ePAQ EEACT was presented as a poster at ISPOR the 16th Annual European Congress November 2-6, 2013 in Dublin, Ireland (Appendix 7), and is currently being written up for publication.

Publications

- **Brennan VK,** Dixon S, Jones G, Radley S, Jacques R, Wood H & Ledger W (2013) An Economic Evaluation Alongside a Clinical Trial (EEACT) in Pelvic Floor Medicine. Value in Health Vol. 16(7) (pp A334-A334) (Poster)

- Jones GL, Radley S, Dixon S, Jacques R, **Brennan V**, Wood H & Ledger W (2013) Evaluating the impact of a 'virtual clinic' on the quality and cost of patient care in urogynaecology: an RCT. BJOG-An International Journal of Obstetrics and gynaecology, , 120, pp 287-288. (Poster)

8.2.5 Incorporating Process Utility into Quality Adjusted Life Years: A Systematic Review of Empirical Studies

Chapter 4 reports a literature review into studies reporting an empirical measure of process utility, which are compatible with the QALY. The literature review was published: “Brennan VK, Dixon S, Incorporating Process Utility into Quality Adjusted

Life Years: A Systematic Review of Empirical Studies, Pharmacoeconomics, 2013 Aug;31(8):677-91". This review included 15 studies which met the inclusion criteria. At the time of publication no existing reviews were identified. The studies were published between 1996 and 2012, and completed in the USA, Australia, Scotland, the UK, Europe and Canada. Studies were identified which explored processes utility associated with treatments, screening procedures, tests and preventative care. A variety of approaches were used to detect and measure process utility, these included SG, TTO, conjoint analysis, wait trade-off, the SF-36 and the EQ-5D. The literature supported the existence of process utility in all indications, although there was considerable variation between estimates. However, due to the range of approaches taken, coupled with the need for further research into the application of the estimates in economic models, it was difficult to know whether the differences were a true reflection of the amount of process utility that enters into an individual's utility function, or whether they are associated with features of the studies methodological design. Despite these considerations, the review indicated that a failure to incorporate process utility into the QALY calculation may mean that health economists are reporting cost-utility analyses that present inaccurate estimates of cost-effectiveness.

This review was accepted and presented as a poster at ISPOR USA in 2012, and published in Pharmacoeconomics in 2013. This led to a letter to the Editor, and the publication of a response. The Pharmacoeconomics review has been cited within more recent studies discussing process utility (Higgins et al. 2014; Poley, 2015; Matza et al, 2014; Swan et al. 2015; Gandjour, 2014; Wisløff et al. 2014).

Publications

- **Brennan VK,** Dixon S. Incorporating process utility into quality adjusted life years: a systematic review of empirical studies. Pharmacoeconomics 2013 Aug;31(8):677-91
- **Brennan VK,** Dixon S. Response to Letter to Editor: Capturing Disutility from Waiting time. April 2014, 32(4)421-422.
- **Brennan VK**, Dixon S, PRM37 A Literature review of Empirical Studies of Process Utility. Value in Health, 15(4)2012. PA165. (Poster)

8.2.6 Developing the health states

Chapter 5 describes the development and results of the process utility valuation study. The foundation of this valuation study was the data collected from within the ePAQ EEACT. The aim was to utilise these data to develop a bolt-on to the SF-6D, an existing multi-attribute preference based measure. Whilst this study is based on the methodology of a previous bolt-on study (Krabbe et al. 1999), the approach taken for the development of the health states was novel for the following reasons:

- The information representing process was one component of the PEQ. This is an existing and established instrument which has undergone psychometric testing for reliability and validity. No other studies have taken this approach.
- This approach allowed the valuation study to be based on valuations of an existing PBM and the bolt-on, as opposed to the use of hypothetical health states.

- This is the first bolt-on study which adds a domain to the SF-6D. Previous studies added domains to the EQ-5D (Krabbe et al. 1999; Yang et al. 2008). The SF-6D is more complex, with more domains, and differing numbers of levels within domains.

8.2.7 Valuing the health states

The methods used for the valuation of health states were also novel:

- The valuation study was completed online, in comparison to Krabbe et al. (1996) who sent out postal questionnaires. This involved working with a software development company to develop online VAS questions and allowed the survey to be undertaken by a larger study sample, more comparable to the general population.
- Using the data already collected from within the ePAQ trial, allowed the study findings to be related back to the clinical trial. This has the advantage that it again reduced the focusing effect reported by Brazier et al. (2011), and also led to an increase in internal validity.
- The derivation of the data collected from within the process utility valuation study was explored further qualitatively, using think aloud techniques. This provides a more in depth exploration of the figures derived from the valuation study than reported by existing studies.
- This study reports cost-effectiveness outcomes both before and after the inclusion of process utility, providing a clear and visual indication of the potential impact of including process. No other study provides a CEAC or cost-

effectiveness plane prior and post the addition of the domain. However, it should be remembered that due to the assumptions made within this study, the revised cost-effectiveness outcomes provide only an indication of the impact of process on cost-effectiveness outcomes.

8.2.8 Think Aloud study into Health Scenarios Combining Health and Process

The process utility valuation study led to a process domain being added to the SF-6D. This resulted in descriptions of health scenarios containing both health and process outcomes. These differ from the usual health states which are valued which contain only health outcomes. The Think Aloud study therefore aimed to explore respondents' reactions to these scenarios which combine both health and process outcomes to determine if this approach is sensible, and to explore the thought processes and considerations respondents make in the valuation of such health states.

This study furthered the current work in the area for the following reasons:

- The think aloud study reported within this thesis is the first think aloud study which focuses on health states which contain not only health outcomes, but also outcomes associated with the processes of receiving care. This expands on the current literature and provides additional insight into the concepts people consider when completing these combined health state valuation studies.

- Current literature suggests we are not truly able to determine whether the information provided by the respondents whilst completing the think aloud interviews is a true reflection of their thoughts Miller & Brewer, (2003). This study aimed to explore this further. Respondents were therefore asked if they felt their responses to the questionnaire would have been the same or different if they were not required thinking aloud when answering the questions. Although some felt the requirement to think aloud did impact on their responses, others felt it had no impact. This question provided insight into the impact that the request to think aloud had on responses.
- A second question was asked to explore this even further. Respondents were asked if they felt their responses would have been the same or different if I had not been sitting in the room during the think aloud activity. Responses were again mixed, although most felt my presence had no impact on their responses.
- This think aloud study also built on existing literature in terms of the analysis approach and reporting. A framework approach was both used and reported in detail. Only one previous study used the Framework analysis approach, and minimal information was provided on the exact steps competed in their analysis (Mulhern et al. 2012). Conversely the methodology reported within this thesis is explicit and transparent. This is important for future qualitative work in health economics as improved reporting of methods, particularly associated with the analysis will lead to more transparent studies whose findings can be generalised more readily.

8.3 Strengths and limitations of the study.

8.3.1 Strengths

One of the main strengths of this study is that it considers process utility from different perspectives and using different techniques. Although not strictly a mixed methods study, as the qualitative and quantitative methods used answer different research questions, this thesis does provide an in depth study into process utility. It contains an EEACT, a systematic literature review into process utility, the development of an online process utility valuation study using novel methods, the valuation of the health and process states, the addition of the findings back into the EEACT, and a think aloud study to gain even deeper insights into the process utility estimates derived. Process utility is therefore thoroughly considered. None of the studies identified in the reviews had explored process utility to this extent.

The study of process utility reported in this thesis combines a foundation of traditional economic and research methods with new and novel approaches to push our knowledge on the topic forward. It has also highlighted issues we need to consider when designing and undertaking health state valuation studies, including the best representation of the domain in terms of content and format, and respondents' interpretations of the health state descriptions.

8.3.2 Limitations

This thesis does have its limitations. The EEACT protocol reported the cost-effectiveness outcomes as a cost per QALY. However, unfortunately the SF-12 was not initially included within the baseline questionnaires which were sent out to

participants. Therefore, the primary economic analysis which was based on participants with complete cost and utility data included data from a small sample with only 27 in the intervention group and 30 the control. Despite this limitation best and current practice was used to account for the handling of missing data.

Although not a choice based valuation study, the approach taken was successful in determining the feasibility and route of further research into process utility, and was a necessary initial step to take. A lot of the work reported within this thesis is exploratory. Therefore, particularly within the process utility valuation study, certain decisions were made to simplify or overcome methodological barriers in order to ascertain the feasibility of undertaking further more detailed analysis, for example, the representation of the severity of the health states, and the reduction in the number of process domain levels from 5 to 3. The findings from the valuation study, and the revised cost-effectiveness estimates incorporating process utility should therefore be interpreted as an indicator of the presence and magnitude of process utility as opposed to absolute values.

Further biases could have been introduced when considering the selection of the PEQ to represent process. The ePAQ EEACT was powered to detect a difference in the PEQ, and not in the CSQ. Therefore the study was more likely to detect significant differences in the PEQ than the CSQ and it was more likely that differences in the CSQ are due to chance. This was not accounted for in the exploratory analysis and PCA. It was more likely therefore that a component of the PEQ (rather than the CSQ) would be selected to represent as process. It may also be the case that the difference in PEQ scores between trial arms which was used to

represent process was amplified, compared to if a CSQ item or domain had been selected. These biases should be considered when reviewing the study results in the context of a feasibility study.

This process utility study is based on an EEACT where no differences in health outcomes were expected, and this approach could be questioned. However, the EEACT methodology was developed for two reasons. Firstly, it was initially considered that there was a possibility that the intervention could speed up referral times and potentially lead to different referral pathways. In which case, the trial-based data collected would have been useful to measure both of these and form the basis of estimating the health and cost impacts. This route was not pursued and it was not deemed relevant for this thesis which benefits from isolating the "process effect". If it was taken further, it would be more realistic to select a trial with health differences also as this is potentially a more realistic and generalisable scenario. Secondly, the original investigator (SD) considered that the realpolitik of the research grant application process meant that even though an EEACT seemed like overkill, it might be criticised if those data were not collected.

There are inherent limitations associated with the qualitative approach. It is likely that to some degree my own experiences and opinions associated with research, health, health care, the NHS, and health care professionals introduce some bias in my work, both in undertaking the think aloud and semi-structured interviews and their interpretation. This may be of particular significance given my background in both health economics and having worked clinically as a physiotherapist for 10 years. This concept of reflexivity whereby my own assumptions and preconceptions

will introduce biases into the research, for example through the questions included in the topic schedule, my interactions with participants and my interpretation of the data will have to a degree have both guided and affected both the study design and findings.

Through writing this thesis, it is also apparent that there has been relatively little progress over time with incorporating process utility into economic evaluations, despite its early advocacy by influential health economists such as Gavin Mooney, whose work into proceduralism and process utility spans over 15 years (Mooney, 1989; Mooney & Blackwell, 2004). This may be for several reasons, both practical and methodological. Practically, it is difficult to develop a standard definition of process utility, particularly as it can potentially include so many aspects of health care, ranging from treatment attributes to patient dignity, and health care infrastructure (Donaldson & Shackley, 1997; Swan & Stainford, 2003; McAlister, 1994), each of which potentially require subjective judgement for their evaluation. The disparate nature of process utility therefore makes it hard to develop standardised and recognised methods for its incorporation into economic evaluations.

A further barrier to its uptake may be due to its relative importance when compared to health comes: from the perspective of both consumers of health care, and the drivers of health economic research. The think aloud study within this thesis began to explore the importance people placed on process versus health outcomes when considering the communication experience. However, this in itself had limitations with respondents being aware of the study objectives. In reality, it

may be that consumers, when ill, and requiring treatment, are primarily concerned with health outcomes, and place minimal value on process. Even more so, the drivers of health economic research (for example, pharmaceutical companies), are likely to be primarily concerned with maximising standard cost-effectiveness outcomes, and the outcomes required by HTA organisations, and be less concerned with additional process benefits which may only provide minute additional benefit at considerable extra expense.

8.4 Areas for future research

As discussed in Chapter 3, the use of ePAQ in advance of a telephone consultation was compared to standard care within STH. Standard care within STH also included the completion of ePAQ online, this is not however standard practice across the UK. Therefore the cost-effectiveness estimates produced are only useful locally to inform resource allocation decisions. Future work should be undertaken in hospitals where ePAQ is not routinely used. This would provide estimates of cost-effectiveness which are more reflective of the UK as a whole, and which would be more useful in informing national policy decisions.

Given the amount of uncertainty introduced by the variation in long term follow-up costs, it may also be sensible to concentrate wholly on costs directly associated with the consultations. Although the ePAQ EEACT incorporated time off work and spent away from usual activities over the 6 month follow-up period, it did not consider personal transportation costs for attending the clinic consultation. These are likely to differ considerably between study arms, both in terms of patients' time

and travel expenses. Their inclusion would provide a more thorough estimate of cost effectiveness.

This work into process utility confirmed that it does exist, and that patients do care about not only health outcomes, but also the processes associated with receiving health care. The process utility valuation study provided a mechanism for capturing the amount of utility people have for these processes, however, the think aloud study raised issues which made us question the validity of utility estimates derived using the health and process scenarios. Therefore further work should be undertaken to explore process utility in general. When developing cost-effectiveness studies, the relevance of process utility should be considered, and if there are likely to be differences in patients preferences for process outcomes, methods should be developed to capture these differences, ideally in a format which would allow them to be incorporated into the QALY. In order for this to occur it is important that health economists included in the early stages of trial design, and that the cost-effectiveness outcomes are not always seen as secondary to clinical outcomes.

If work on process utility is taken forward, further consideration should also be given to its application in economic models, particularly in relation to the duration of the effect. Within this study, it was assumed that the utility gained or lost from the communication experience lasted for 6 months, and had the same duration as the non-process domains. Although some may agree with this, for others the duration of the impact may have been more or less. Further work should therefore be undertaken in this area, for example into whether processes for health and

process outcomes have the same duration, and whether there are methods which could apply process utility decrements which account for the duration of the process effect.

This process utility analysis within this thesis is based only on the communication domain from the PEQ. However, as outlined in Chapter 2, the term "process utility" encapsulates much more than purely communication. Therefore, when interpreting these results, it should be clear that the study suggests respondents do have a preference for the communication experience gained from either a telephone or face-to-face consultation, and that it is possible to capture this preference within a format which is compatible with the QALY. This study therefore, only looks at one type of process utility (communication). Although it is unlikely there could be one "process domain" developed which could capture respondent's preference for process as a whole, as sub-categories of process utility differ so substantially (e.g. treatment attributes, dignity, communication) it may be that, for example a domain could be developed which considers "patient satisfaction" as a whole, and such a domain may be compatible across different types of process? Therefore, for example, rather than concentrating specifically on the communication domain of the PEQ a domain may be based on an entire "patient experience questionnaire". This study has therefore brought together disparate literatures, including patient experience, patient satisfaction, and cost-utility analyses. Furthering the work, towards the development of a generic satisfaction bolt-on would be both interesting and could result in beneficial advances in cost-utility analyses.

The think aloud study provided valuable insights into health state valuation studies as a whole, and the use of health scenarios which include both health and process outcomes. However, as identified by previous work, we are still unable to ascertain whether the information respondents provided whilst thinking aloud was truly reflective of their thoughts, or whether the process of thinking aloud affected this. Although this study aimed to explore this further with the additional questions on the impact of the think aloud technique this is an area which should be studied further, whether this could be by having a respondent alone in a room completing a questionnaire, but still being recorded, or whether alternative approaches could be developed, which also allowed for prompting when respondents failed to think aloud. Advances in the Think Aloud approach would provide valuable information, not only for this study on process and health domains, but also for future studies into health state valuations used for economic evaluations which currently contain only health outcomes. The more we can understand about the processes respondents go through when valuing health states, the more we can refine our methodologies to reflect more closely human behaviour and move towards providing more accurate estimates of cost-effectiveness. The approach is not exclusively relevant to valuation studies, but in fact should be considered for any "task-based" activities. The ability to gain insight into respondents thought processes when completing tasks is invaluable. This could be for example, the use of SG or TTO approaches, or the thoughts a person has when completing a WTP questionnaire. Utilising the think aloud approach would provide an additional layer of knowledge above just the questionnaire responses themselves and the use of this qualitative approach to understand more fully the processes which are used to

obtain the quantitative information inputted into economic evaluations could only serve to enhance our application of these methods within economic evaluations.

In addition to this, further work should be undertaken on the framework approach which was utilised within this study to analyse the think aloud transcripts. One of the main benefits provided by this approach is the transparency of the findings, displayed in a matrix, allowing readers to view the information, make more informed decisions on the generalisability of study findings and their application to different economic evaluations. The think aloud approach seems to be a valuable tool which provides us with a more in depth insight into the way people value health states. Therefore, it could only benefit from exploring this further, both in terms of the impact thinking aloud has and responses, and ways to address any issues which may be identified.

8.5 Conclusion

With the aim of improving the relevance of the cost-effectiveness outcomes from the ePAQ EEACT this thesis presents a study of process utility. It highlights that, although studies of cost-effectiveness routinely capture patients preferences associated with health outcomes, they do not always capture differences associated with the processes of receiving care. The ePAQ EEACT has highlighted that in some instances these are the most important outcomes to consider.

The findings from this study therefore question the assumption of consequentialism in determining the QALY (that consumers of health care gain no benefit from its

consumption, but solely from the health outcomes it generates), and indicate that important elements of a patient's health care utility function are being ignored. The revised cost-effectiveness estimates from the ePAQ trial which moved the intervention from a position where it was dominated by the control (North West quadrant: where the intervention is more costly and less effective), to a position where the intervention is more effective and more costly than the control (North East quadrant), and to a position where it also falls within the willingness to pay threshold of £20,000/QALY indicates that the inclusion of process utility can have an impact on cost-effectiveness results, and therefore important implications for resource allocation and policy decisions.

8.6 References

- Brennan, VK. and Dixon, S. (2013) Incorporating Process Utility into Quality Adjusted Life Years: A Systematic Review of Empirical Studies, *Pharmacoeconomic*s, 31(8), 677-91

- Brennan, VK. Dixon, S. Jones, G. Radley, S. Jacques, R. Wood, H. Ledger, W. (2013) An Economic Evaluation Alongside a Clinical Trial (EEACT) in Pelvic Floor Medicine. *Value in Health*, 16(7) (pp A334-A334) (Poster)

- Donaldson, C. Shackley, P. (1997) Does "process utility" exist? A case study of willingness to pay for laproscopic cholcystectomy. *Soc Sci Med*, 44, 699-707.

- Drummond, MF. and Stoddart, GL. (1984) Economic analysis and clinical trials. *Controlled clinical trials* 5,115-128.

- Friedberg, M. Saffran, B. Stinson, TJ. Nelson, W. Bennett, CL. (1999) Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA,* 282(15).

- Gandjour, A. (2014) Capturing Disutility from Waiting Time. *PharmacoEconomics*, 32(4), 423.

- Higgins, A. Barnett, J. Meads, C. Singh, J. Longworth, L. (2014) Does Convenience Matter in Health Care Delivery? A Systematic Review of Convenience-Based Aspects of Process Utility. *Value in Health*, 17,877-887.

- Jones, GL. Radley, S. Dixon, S. Jacques, R. Brennan, V. Wood, H. Ledger, W. (2013) Evaluating the impact of a 'virtual clinic' on the quality and cost of patient care in urogynaecology: an RCT. *BJOG-An International Journal of Obstetrics and gynaecology*, 120,287-288. (Poster)

- Krabbe, P. Stouthard, M. Essink-Bot, M. Bonsel, G. (1999) The effect of adding a cognitive dimension to the EuroQol Multiattribute Health-Status Classification System. J clin Epidemiol, 52(4), 293-301.

- Matza, LS. Sapra, SJ. Dillon, JF. Kalsekar, A. Davies, EW. Devine, MK. Jordan, JB. Landrian, AS. Feeny, DH. (2014) Health state utilities associated with attributes of treatments for hepatitis C. *The European Journal of Health Economics,* Dec, 7.

- McAlister D. (1994) Putting health economics into quality. Public Money & Management, 14(2),15-22

- Mooney, G. (1989) QALYs: are they enough? A health economist's perspective. Journal of medical ethics. 15,148-152.

- Mooney, GH. Blackwell, SH. (2004) Whose health service is it anyway? Community values in healthcare. *Med J Aust.* 180(2).76-8.

- Mulhern, B. Tsuchiya, A. Brazier, J. Rowen, D. (2012) How do respondents perceive health state valuation exercises? A think aloud study investigating time trade off and discrete choice experiments. Paper for presentation at EuroQol Plenary Rotterdam.

- Poley MJ (2015) Nutrition and health technology assessment: when two worlds meet. Front. *Pharmacol*. 6:232.DOI: 10.3389/fphar.2015.00232.

- Swan, JS. Kong, CY. Hur, C. Halpern, EF. Itauma, O. Williams, O. Motazedi,T. Lee, JM. (2015) Comparing Morbidities of Testing With a New Index: Screening Colonoscopy Versus Core-Needle Breast Biopsy, *Journal of the American College of Radiology*, 12(3),295.

- Swan, J. Sainfort, F. (2003) Process Utility for Imaging in cerebrovascular disease. *Acad Radiol,* 10(3), 266-74.Wisløff, T. Hagen, G. Hamidi, V. Movik, E. Klemp, M. Olsen, JA. (2014) Estimating QALY gains in applied studies: a review of cost-utility analyses published in 2010. *Pharmacoeconomics*,32(4),367-75.

- Yang, Y. Brazier, J. Tsuchiya, A. (2008) The effect of adding a 'sleep' dimension to Q-5D. Health Economists' Group meeting, January

Appendix 1: Literature review into limitations of EEACTs. List of included studies.

1. Baltussen, R. Ament, A. Leidl, R. (1996) Making cost assessments based on RCTs more useful to decision-makers. *Health Policy,* 37(3), 163-83.

2. Baltussen, R. Leidl, R. Ament, A. (1999) Real world designs in economic evaluation. Bridging the gap between clinical research and policy-making. *Pharmacoeconomics*, 16(5 Pt 1), 449-58.

3. Bennett, CL. (1995) Economic Analysis during Phase III clinical trials: who, what, when, where and why? *Oncology*. November 01.

4. Bennett, CL. Smith, TJ. George, SL. Hillner, BE. Fleishman, S. Niell, HB. (1995) Free-riding and the prisoner's dilemma: problems in funding economic analyses of phase III cancer clinical trials. *J Clin Oncol*, 13(9), 2457-63.

5. Bombardier, C. Maetzel, A. (1999) Pharmacoeconomic evaluation of new treatments: efficacy versus effectiveness studies? *Ann Rheum Dis*, 58(Suppl 1),I82-5.

6. Bonsel, GJ. Rutten, FF. Uyl-de Groot CA. Economic evaluation alongside cancer trials: methodological and practical aspects. Eur J Cancer. 1993;29A Suppl 7:S10-4.

7. Brennan, A. Akehurst, R. (2000) Modelling in Health Economic Evaluation. What is its place? What is its value? *Pharmacoeconomics* 17(5) 445-459.

8. Briggs, A. Tambour, M. (2001) The design and analysis of stochastic cost-effectiveness studies for the evaluation of health care interventions. *Drug Inf J*,35:,455–68.

9. Briggs, A. Joint MRC/Southeast Region training fellow (2000) Economic evaluation and clinical trials: size matters : The need for greater power in cost analyses poses an ethical dilemma. *BMJ*. 321(7273):1362-1363.

10. Byford, S. Palmer, S. (1998) Common errors and controversies in pharmacoeconomic analyses. *Pharmacoeconomics* 13 (6), 659-66

11. Cook, J. Drummond, M. Heyse, JF. Economic endpoints in clinical trials. *Stat Methods Med Res*. 2004 Apr;13(2):157-76.

12. Cookson, R. Hutton, J. (2003) Regulating the economic evaluation of pharmaceuticals and medical devices: a European perspective. *Health Policy* 63(2003) 167-178.

13. Cox, HL. Laupland, KB. Manns, BJ.(2006) Economic evaluation in critical care medicine (2006). *Journal of Critical Care*. 21(2),117-124.

14. Coyle, D. Davies, L. Drummond, MF. (1998) Trials and tribulations. Emerging issues in designing economic evaluations alongside clinical

trials. *International Journal of Technology Assessment in Health Care*, 14 (1) 135-144

15. Coyle, D. Lee, KM. (1998) The problem of protocol driven costs in pharmacoeconomic analysis. *Pharmacoeconomics.* 14(4),357-363.

16. De Maio, E. Perrone, F. Bisagni, G; et al. (2006) The new drugs advent: clinical or economic outcomes? *Ann Oncol*. 17 (Suppl 2) :ii88-90.

17. Dixon, S. Deverill, M. Gannon, M. Brazier, J. Haggard, M.(1999) Improving the validity of economic evaluations alongside controlled trials *Journal of Health Services and Research Policy*, 4(3),161-163

18. Dowie, J. (1997) Clinical Trials and economic evaluations? No there are only evaluations. *Health Economics* 6 87-89.

19. Drummond, M. Jefferson, T. (1996) Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ* 313:275-283.

20. Drummond, M. Sculpher, M. (2005) Common methodological flaws in economic evaluations Source: *Medical Care* 43 (7),5-14 Suppl.

21. Drummond, M. (1995) Economic analysis alongside clinical trials: problems and potential, *Journal of Rheumatology* 22 (7),1403-1407

22. Drummond MF, Davies L. (1991) Economic analysis alongside clinical trials. Revisiting the methodological issues. *Int J Technol Assess Health Care,* 7(4):561-73.

23. Drummond, MF. Stoddart, GL. (1984) Economic analysis and clinical trials. *Controlled clinical trials,* 5:115-128.

24. Gheorghe, A. Kyte; D. Calvert, M. (2014) The need for increased harmonisation of clinical trials and economic evaluations Expert Rev Pharmacoecon Outcomes Res, 14(2),171-3.

25. Haycox, A. Drummond, M. Walley, T. (1997) Pharmacoeconomics: integrating economic evaluation into clinical trials. *Pharmacoeconomics* ,43(6):559-62

26. Hlatky, MA. Owens, DK. Sanders, GD. (2006) Cost-effectiveness as an outcome in randomized clinical trials. *Clin Trials*. 3(6),543-51.

27. Imai, K. Zhang, P (2005) Integrating economic analysis into clinical trials. *LANCET* , 365(9473), 1749-1750

28. Jackson, JD. (2010) Economic evaluations alongside effectiveness trials. *Value Health*, 13(8):865-6.

29. Thorn, JC.; Noble; SM. Hollingworth, W. Methodological developments in randomized controlled trial-based economic evaluations *Expert Rev Pharmaoecon Outcomes Res,* 14(6),843-56.

30. Khan Z, Miller D (1999) Modelling Economic Evaluations of pharmaceuticals: Manipulation of valuable tool? *Clinical therapeutics* 21(5).

31. Marshall, DA. Hux, M. (2009) Design and analysis issues for economic analysis alongside clinical trials *Med Care.* 47(7 Suppl 1):S14-20.

32. O'Brien, B. (1996) Economic Evaluation of Pharmaceuticals. Frankenstein's monster or Vampire of trials? *Medical Care*. 34(12)DS99-DS108 Supplement

33. O'Sullivan, A. Thompson, D. Drummond, M. (2005) Collection of health economic data alongside clinical trials: is there a future for piggyback evaluations? *Value in Health* 8(1).

34. Ramsey, SD. McIntosh, M. Sullivan, SD. (2001) Design issues for conducting cost-effectiveness analyses alongside clinical trials. *Annu Rev Public Health*. 22:129-41

35. Ramsey, SD. Willke, RJ. Glick, H. Reed, SD. Augustovski, F. Jonsson, B. Briggs, A. Sullivan, SD.(2015) Cost-effectiveness analysis

alongside clinical trials II-An ISPOR Good Research Practices Task Force report. *Value Health*. 18(2),161-72.

36. Revicki, DA; Frank, L (1999) Pharmacoeconomic evaluation in the real world - Effectiveness versus efficacy studies *Pharmacoeconomics*. 15(5),423-34.

37. Rittenhouse B (1996) Uses of models in economic evaluations of medicines and other health technologies. Office of Health Economics, BSC Print, London.

38. Sculpher M (2015) Clinical Trials Provide Essential Evidence, but Rarely Offer a Vehicle for Cost-Effectiveness Analysis. *Value in Health,* 18,141-142.

39. Sculpher, MJ. Claxton, K. Drummond, M. et al. (2006) Whither trial-based economic evaluation for health care decision making? *Health Economics*, 15 (7),677-687

40. Sheldon, T. (1996) Problems of using modelling in the economic evaluation of health care. *Health Economics* 5:1-11.

41. Friedberg, M. Saffran, B. Stinson, TJ. et al. (1999) Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA* 282(15).

Appendix 2: ePAQ EEACT Study Documents

This appendix presents the ePAQ study paperwork. This includes the study protocol, a flow chart of the ePAQ study, the proforma for the expression of interest, the ePAQ patient information sheet, a patient consent form, the demographics proforma and a copy of the telephone consultation record.

These documents were originally prepared by Dr Georgina Jones of the University of Sheffield, and submitted to ethics for approval. They were then updated by Victoria Brennan in a substantial amendment to the North Sheffield Ethics Committee for the inclusion of the SF-12. All amendments made by Victoria remain highlighted.

Collaborating Organisations:

STHFT Department of Obstetrics & Gynaecology

&

University of Sheffield

Health Services Research Section, ScHARR

Assessing the impact of using ePAQ on quality and cost of patient care in urogynaecology:

an RCT (STH14733)

Study Protocol

Proposed Duration	Proposed Starting Date
24 months	01 December 2007

Members of the Research Team

Mr Stephen Radley

Job Title: Consultant Gynaecologist

Organisation: Sheffield Teaching Hospitals

Department: Department of Obstetrics and Gynaecology

Full Address: Level 4, Jessop Wing Hospital

Phone Number: 0114 2268574

Fax Number: 0114 2268165

Email Address: stephen.radley@sth.nhs.uk

Dr Georgina Jones

Job Title: Senior Lecturer

Organisation: University of Sheffield

Department: Health Services Research Section, ScHARR

Full Address: Regent Court, 30 Regent Street, Sheffield, S1 4DA

Phone Number: 0114 2220806/ 0114 2268515

Fax Number: 0114 2422136

Email Address: g.l.jones@sheffield.ac.uk

Dr Swati Jha

Post: Subspecialty Trainee in Urogynaecology

Organisation: Sheffield Teaching Hospitals NHS Trust

Department: Department of Obstetrics and Gynaecology

Full Address: Level 4, Jessop Wing Hospital

Phone Number: 0114 2268167

Fax Number: 0114 2268165

Email Address: Swati.Jha@sth.nhs.uk

Dr Simon Dixon

Job Title: Senior Lecturer

Organisation: University of Sheffield

Department: Health Services Research Section, ScHARR

Full Address: The University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA

Phone Number: 0114 222 0724

Email Address: s.dixon@sheffield.ac.uk

Dr Stephen Walters

Job Title: Senior Lecturer

Organisation: University of Sheffield

Department: Health Services Research Section, ScHARR

Full Address: The University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA

Phone Number: 0114 222 0730

Email Address: s.j.walters@sheffield.ac.uk

Professor William L. Ledger

Job Title: Professor of Obstetrics and Gynaecology & Honorary Consultant in Obstetrics and Gynaecology.

Organisation: University of Sheffield

Department: Academic Unit of Reproductive & Developmental Medicine

Full Address: Room 5, Level 4 Jessop Wing Tree Root Walk Sheffield S10 2SF

Phone Number: 0114 226 8317

Email Address: w.ledger@shef.ac.uk

Title

Assessing the impact of using ePAQ on quality and cost of patient care in urogynaecology: an RCT

Background to the Research

What is ePAQ?

ePAQ is a web-based interactive questionnaire (electronic pelvic floor assessment questionnaire) that provides a detailed evaluation of a woman's pelvic floor symptoms and their impact on her quality of life. Touch-screen technology has been combined with computer programming (Dot Net & SQL Server) and existing paper-based instruments to produce an interactive user-friendly system that facilitates immediate data entry, analysis and storage. The instrument has evolved during several years' collaborative work with colleagues in a range of disciplines, including Gynaecology, Medical Physics, Colorectal Surgery, Urology and General Practice, as well as involving a women's health user-group and regular surveys of patients' views.[1] Research in primary and secondary care has established the psychometric properties of the instrument.[2] The system has won a National award for innovation (HITEA) and was cited in Lord Warner's Annual Report on NHS Innovation.[3] EPAQ Systems Ltd is one of the first NHS spinout technology companies to be established in the UK (created jointly with Sheffield Teaching Hospitals NHS Trust in November 2006).

More recent development work has resulted in the creation of a website, where subjects can securely and anonymously access and use the ePAQ system via the Internet. This potentially allows subjects to use system in advance of clinic appointments (as opposed to the current practice of their completing it on arrival in clinic, immediately prior to their appointment). This study aims to evaluate the impact that using ePAQ in this way may have on the efficiency and quality of patient care. Initial survey data suggest that many women attending urogynaecology clinics do have access to the Internet and would be willing to use the system on-line & in advance of clinic appointments.[4] The facility is also available in the Gynaecology Unit for women without Internet access to attend specifically for supervised use of the system. The in-depth assessment provided by the ePAQ may then be used to support an initial telephone consultation with a clinician, following which patients may be directed to the most appropriate clinic, as well as being provided with information and advice and have some forms of treatment initiated (such as behavioural therapy). The proposed randomised study aims to compare outcomes in

women who use the ePAQ system in this way with women who undergo standard care in the Urogynaecology Unit.

The major advantage of electronic systems, when compared with paper questionnaires, relate to the practicalities of clinical data capture. Electronic questionnaires are comparable in terms of reliability and can be superior in terms of efficiency and response rate. In addition, electronic touch-screen questionnaires have been shown to be acceptable to patients, regardless of their age or educational background. Data quality is high, even in disadvantaged and technophobic subjects and cost analysis has shown potential economic advantages.[5,6,7,8,9] Pelvic floor disorders in women, such as incontinence and prolapse, share common aetiologies and commonly coexist. Estimates vary, though it is estimated that approximately 20% of adult women experience regular urinary incontinence, 5% have some incontinence of faeces and 11% suffer with prolapse.[10,11,12] However, despite better understanding of these conditions, bowel, bladder, vaginal and sexual dysfunction remain understandably taboo subjects and many women still regard them as inevitable consequences of childbirth and ageing. The personal cost to individuals is high, however simple and effective treatments are increasingly available. Many disorders respond well to behavioural, physical or medical therapy and such conservative treatments are generally recommended prior to consideration of invasive investigations or surgery.[13] In clinical practice, clinical assessment is central to diagnosis and management and the restoration of function, with a view to improving quality of life, is the principal aim of treatment. It is well recognised that clinical interview data may be unreliable, being based on clinicians' rather than patients' views of their condition. It seems appropriate therefore, to seek ways of improving clinical assessment in order to enhance the quality of care and reliably measure outcome. Many women with pelvic floor disorders are managed in the community by GPs, nurses, physiotherapists or continence nurse advisors. In secondary and tertiary care, urologists, gynaecologists, colorectal surgeons and geriatricians are likely to be involved. However, at all levels, inconsistencies in clinical assessments represent an impediment to effective communication and the multidisciplinary approach advocated by the DoH, who in 2000 recommended the following:

- Full assessment leading to first line treatment in the primary care setting, with treatment & management plans agreed with individual patients.

- The provision of an integrated continence service, bringing together agreed protocols and procedures for primary, secondary and tertiary care.

• A comprehensive continence service, at home and in homes, bringing together all relevant health disciplines.

The further development of ePAQ has the potential to substantially augment this process by using on-line assessments to support patients' initial management and triage. Using ePAQ in advance of clinic appointments may provide patients and clinicians with prompt and valuable information and assist in directing patients to the most appropriate clinics in primary and secondary care. However, such a development warrants scrutiny in terms of patient experience and cost.

3. Specific Research Question(s) / Hypotheses

The ePAQ is now an established part of standard care & is in routine use in the Sheffield Urogynaecology Unit; All new patients are given the opportunity to complete the questionnaire on arrival in clinic and their printed results are then used to inform the subsequent clincial consultation. The aims of this research are to measure the impact that using ePAQ *in advance* of clinic appointments in combination with a telephone consultation has on patient care. Using an RCT methodology we will specifically evaluate its impact on the following:

i) Patients' experience of the clinical episode

Outcome measures: Time to first assessment, Time to initiation of treatment, Patient Experience Questionnaire

ii) Empowerment of patients in relation to their condition and healthcare that they receive

Outcome measure: Modified CSQ-8 (Client Satisfaction Questionnaire), QQ-10

iii) Cost, efficiency and quality of life.

Outcome measures: Number & type of clinic attendances & consultations, Costs of prescriptions & investigations, SF-12.

4. Proposed research method.

a) Study design

Randomised controlled trial. All women who consent to participate will be randomised to one of 2 groups:

1) ePAQ + telephone consultation

or

2) Usual care

b) Description of the methodology

This research will follow a quantitative methodology.

c) Time frame

24 months

Recruitment = 12 months. Follow-up = 6 months. Data analysis & preparation of results for publication = 6 months

d) Sample

All women referred to STH urogynaecology services, aged > 18 and able to read and understand English will be eligible to enter the study. Potential recruits to this study will be identified by review of referral letters received, from GPs or other clinicians, by the

consultant to whom the patient has been referred (SC Radley or AG Farkas). It is envisaged that a minimum of 6 new patients per week will be recruited to this study.

e) Sample size calculation

The primary outcome for the purposes of sample size determination is the mean outcome scale score on the Patient Experience Questionnaire (PEQ) completed following the first consultation (Steine et al, 2001). 304 patients will be recruited (assuming a 20% dropout rate) to provide the 121 women needed in each group to detect significant differences between women who use ePAQ in advance of clinic appointments and women undergoing usual care (see Table 1 below).

Table 1.

Sample size for PEQ outcome scale at 1% and 5% significance levels and various standard deviations

	1	2	3	4	5	6
Test significance level	0.01	0.01	0.01	0.05	0.05	0.05
1 or 2 sided test?	2	2	2	2	2	2
Difference in means	0.5	0.5	0.5	0.5	0.5	0.5
Common standard deviation	1.2	1.1	1.0	1.2	1.1	1.0
Effect size	0.42	0.46	0.50	0.42	0.46	0.50
Power (%)	90	90	90	80	80	80
n per group	174	146	121	92	77	64
Total N	348	292	242	184	154	128

Total N (with 20% dropout)	436	366	304	230	194	160

Steine (et al 2001) reported a mean score of 2.9 (SD 1.0) for the outcome scale of the PEQ and mean score of 5.0 (SD 1.2) for the emotion scale. If we assume an SD of 1.0 for the outcome scale and that a mean difference of 0.5 or more points between the intervention and control groups is of both clinical and practical importance, then to achieve a 90% power for demonstarting this mean difference as being statistically significant at the 1% (two sided) level will require 121 women per group (242 in total). Assumingthat 20% of patients do not return completed PEQs, then it will be necessary to recruit and randomise 304 patients (152 per group).

Recruitment (see flow chart: appx I)

Typically women are referred (by letter) by General Practitioners, other Hospital Specialists or Primary Care Continence Services. All referral letters will first be reviewed by the consultant (AGF or SCR) to whom the patient has been refferred. The consultant will then decide whether or not patients is suitable for study entry. Patients considered suitable by their consultant will be contacted by telephone by the Research Nurse to discuss the study (appx II). In order to prevent any significant delays in potential study patients being offered clinic appointments, patients who are not contactable by telephone within 48 hours of the reciept of the referral letter will be excluded from the study and be sent a routine outpatient clinic appointment. Patients who express no interest in participating in the study will also be sent a routine outpatient clinic appointment. Patients who express an interest in participating in the study will be sent an information leaflet about the study and a consent form (appendix III) with SAE. Women who do not respond to this within 1 week will be contacted by telephone to assess whether they wish to participate in the study. Those women who cannot be contacted at this point, or who decline study entry will be be sent a routine outpatient clinic appointment. On receipt of signed consent forms subjetcs will be entered into the study and randomised. At this point, patients randomised to standard care will be posted clinic appointments with AGF or SCR. Women randomised to ePAQ will be posted an information letter and epaq-online information leaflet inviting them to complete the questionnaire on-line and to telephone Gynae Appointments in order to arrange their telephone consultation (with AGF or SCR). Subjects in this group who are unable to

complete epaq-online will use this telephone call with gynaecology appointments to arrange a separate visit to the hopsital to complete the questionnaire.

f) Randomisation

Randomisation will occur on receipt of signed consent forms. Allocation will be through stratified block randomisation. SW will generate a randomisation schedule, for these strata, using the STATA software. This will be held remotely by GLJ (who will not be directly involved in recruiting patients to the study). By referring to the list, she will allocate patients to either the intervention group (ePAQ + telephone consultation) or the control group (usual care). The randomised allocation will then be passed on to the research nurse who will then send the relevant information packs to each participant, relating to the arm of the study to which they have been randomised.

Intervention Group: ePAQ clinic

All women randomised to this arm of the study will be enetered onto the hospital PAS system under the ePAQ clinic. They will be invited by letter to complete epaq-online (and if necessary supervised ePAQ completion as a separate visit) and asked to contact Gynaecology Appointments in order to arrange their telephone consultation with AGF or SCR:

1. Letter inviting patients to complete epaq-online (including epaq-online information leaflet) & to telephone Gynaecology Appointments in order to book their telephone consultation.
2. Patient telephones Gynaecology Appointments when clerk will:
    1. Ascertain whether patient has already succesfully completed epaq-online

       If not: Agree & send a date, time & place to attend for supervised completion of ePAQ in hospital.

    2. Agree & send (to all patients) a date, time & number for their telephone consultation
3. Scheduled telephone consultation with AGF or SCR

   (supported by ePAQ report, referral letter and hospital casenotes)

4. Letter dictated to patient (including appropriate patient information leaflets & follow-up appointment)
5. Letter written to GP & referring clinician (including copy of above letter to patient)

The epaq-online information leaflet includes details of the web address for epaq-online (www.epaq-online.co.uk) with information explaining how to log-on and use the website. On completion of the questionnaire, the patient has the option of printing out a summary of her questionnaire data, which will include a summary of symptom scores in each area (urinary, bowel, vaginal and sexual). The patient is asked to forward her username and questionnaire ID to an STH email address (epaq@sth.hns.uk). Data are encrypted and anonymised (in line with good data protection practice) and can only be personalised to an individual patient by an approved clinician in receipt of the patient's username & questionnaire number, provided by the patient themselves. The clinician will use these results along with the patient's own casenotes and original referral letter to support the subsequent telephone consultation.

Non-intervention (control) group: Usual care

All women randomised to this arm of the study will be given an appointment to attend the urogynaecology clinic. 'Usual care' currently includes the option of completing the ePAQ on arrival in clinic, immediately prior to the clinical consultation. As is currently standard practice, the results of the ePAQ will be used to inform & support clinical assessment, however, as the ePAQ is used *immediately* prior to the clinical consultation, patients cannot be triaged or receive any additional information on the basis of their questionnaire results.

Outcome Measures

In order to achieve uniformity of approach (between ePAQ and control groups) outcome questionnaires (PEQ, CSQ-8, QQ-10 & SF-12) will be posted to patients *immediately* after their first consultation. Patients will be asked to complete and return these by post.

Primary Outcome Measure

Patient Experience Questionnaire

The primary outcome for the purposes of sample size determination is the mean outcome scale score on the Patient Experience Questionnaire (PEQ) (Steine et al, 2001) (Appendix 6). The PEQ will be used to evaluate whether using ePAQ in combination with a telephone consultation results in any difference in the patient experience when compared with standard care. In order to evaluate the impact of using ePAQ in this way, we ask all patients (in both arms of the study) to complete the PEQ after their first clinical consultation.. The PEQ was developed and validated specifically to measure patients' experience of interaction, emotion and consultation outcome. It contains 16 items in four dimensions, including 1) communication, 2) emotions 3) short-term outcome and 4) barriers. Three scales are scored from 1 to 5, and the emotion scale runs from 1 to 7. A high score represents a good communication experience, positive emotions, positive consultation outcome and a lack of communication barriers.

Secondary Outcome Measures

1. Number and type of referrals

The total number and individual type of referrals made will be used to evaluate the impact of ePAQ on the referral patterns and clinical visits made within the continence care pathway. At 6-month follow-up we will measure the number and type of referrals made on behalf of that patient; i.e. primary care (community nursing, community physiotherapy, GP, district nurse, practice nurse) and secondary care (gynaecology, urology, colorectal surgery, physiotherapy). These data will be collected in both groups using a standard proforma (appx VII).

2. Number and type of prescriptions

The number and type of prescriptions made will be collected to evaluate whether the use of ePAQ in advance of clinic appointments has any influence on the number of prescriptions made within the continence care pathway. These data will be collected at the 6-month follow-up on a proforma (appx VII) and supported by looking through the appropriate medical records. The prescribing of first line treatment will be categorised as (a) drug treatment (b) physiotherapy (c) behavioural therapy and (d) surgery.

4. Modified Client Satisfaction Questionnaire

A modified version of the CSQ-8 will be administered (Howie et al, 1997; 1998) (Appendix 7). This validated instrument contains 8 items measuring the patient's satisfaction following a consultation asking 1) how they rate the service, 2) did they get the service they wanted, 3) to what extent the service me their needs, 4) would they recommend the service to their friends, 5) how satisfied they were with the amount of help received, 6) if the services received helped then to deal more effectively with their problems, 7) overall satisfaction with the service and 8) if they would return to the services if they required help again.

5. Cost effectiveness

The economic analysis will estimate the NHS costs of providing assessment and care up to six months in both arms of the study, and cost effectiveness by an incremental cost per quality-adjusted life year (QALY). These costs will include the costs of using ePAQ + telephone consultations, referrals, prescriptions (drugs / appliances / pads) and treatments. A broader societal costing will also be undertaken to consider costs to the patients and the economy in terms of private expenditure and time taken off work, respectively. These additional data will be in the standard proforma posted to both groups of patients at 6 months (appx VII).

Data on the use of the health services will be collected for each patient from records accessed from Trust and Primary Care practices. Data on private and societal costs will be gathered using patient-completed questionnaires at 6 months (appendix VII). Such patient-completed questionnaires are used in nearly all economic evaluations. Examples of questions include; "In the last 6 months, have you spent any money on anything because of bladder, bowel or vaginal problems?", "If yes, what did you spend it on?" "How much have you spent on each?", and "In the last 6 months, how many days have you needed to take off work because of bladder, bowel or vaginal problems and their treatment / care?" Costs will then be calculated using unit costs from routine sources (Curtis & Netten 2004, NHS Reference Costs and the British National Formulary). The costs associated with using ePAQ + telephone consulations compared with usual care will be estimated through a micro-costing study conducted during the first full 2-week period following recruitment of the 200th patient to the study.

The quality of life of patients will be explored using the SF-12 at 2 time points within the study. Data will be collected prior to the patients initial consultation to gain a baseline

quality of life estimate. This will be sent to consented patients with their randomisation details. The second measurement will be taken after 6 months and the questionnaire will be included with the 6 month outcome measurement questionnaire. This will highlight any longer term differences in quality of life. This quality of life data, combined with both the clinical outcomes and costs, will be used to calculate an incremental cost per QALY.

h) Planned Analysis:

All ePAQ data are held on a secure password protected database. All data will be anonymised and transferred to SPSS for Windows (version 14.0), which will be used to carry out all statistical analyses. As the trial is a parallel group RCT, data will be reported and presented according to the revised CONSORT statement (Moher et al, 2001). The statistical analyses will be performed on an intention-to-treat (ITT) basis. All statistical exploratory tests will be two-tailed with $P < 0.05$. Baseline demographic data will be assessed for comparability between the treatment groups. Other secondary outcomes are to be assessed at baseline (randomisation) and at 6 months post randomisation. In the event of differences between the ePAQ and usual care groups with respect to baseline demographic, physical, and health-related quality of life measurements, multiple regression or analysis of covariance (ANCOVA) will be used to adjust the treatment effect for these variables. The ordinary least squares adjusted regression coefficient estimate for the treatment group parameter along with its 95% confidence interval (CI) will then be reported.

1. Analysis of the PEQ

The primary efficacy response variable is the outcome scale of the PEQ. The PEQ is to be posted to the subject immediately after the consultation with the clinician. A two independent samples t-test will be used to compare mean PEQ outcome scale scores between groups (ePAQ and usual care) in this parameter. A 95% confidence interval (CI) for the mean difference in this parameter between the ePAQ and usual care groups will also be calculated. Secondary efficacy outcomes are the other four scales of the PEQ: communication; emotion; barriers and auxiliary staff. These will be analysed in a similar way to the outcome scale.

2. Analysis of referral patterns

At the 6 month follow-up we will measure (1) the number of referrals made for that patient and (2) the type of referral i.e. primary care (community nursing, community

physiotherapy, GP, district nurse, practice nurse) and secondary care (gynaecology, urology, colorectal surgery, physiotherapy). The number of referrals over the 6-month follow-up will be described and tabulated. A two independent samples t-test will be used to compare mean number of referrals between the groups (ePAQ and usual care) in this parameter. A 95% confidence interval (CI) for the mean difference in this parameter between the ePAQ and usual care groups will also be calculated. The type of referral will be described and tabulated for each group. A chi-squared test will be used to compare referral patterns between the 2 groups (ePAQ and usual care).

3. Analysis of the number of consultations

The number and type of consultations over the 6-month follow-up period will be described and tabulated. An independent samples t-test will be used to compare mean number of consultations between the 2 groups (ePAQ in advance of clinic appointments vs usual care). A 95% confidence interval (CI) for the mean difference in this parameter will also be calculated.

4. Analysis of the number and type of prescriptions

Prescribing patterns (including number and type) over the 6-month follow-up period for each group will be described and tabulated. A two independent samples t-test will be used to compare mean numbers of prescriptions between the groups (ePAQ and usual care) for this parameter. A 95% confidence interval (CI) for the mean difference in this parameter between the ePAQ in advance of clinic appointments and usual care groups will also be calculated. A chi-squared test will be used to compare prescribing patterns between the groups (ePAQ and usual care).

5. Analysis of cost-effectiveness

The incremental cost-effectiveness ratio constructed from the cost and QALY data will be presented along with its associated cost-effectiveness acceptability curve. The SF-12 questionnaire will be used to obtain a quality of life value which will be used for calculate QALY's for each treatment arm. This will be administered at 3 time points i) prior to the patients consultation (sent to consented patients with randomisation details) ii) after the consultation iii) After 6 months with the 6 month outcome assessment questionnaires. Separate analyses will be undertaken for NHS costs and the broader

societal costs. One-way sensitivity analysis will be undertaken to assess the robustness of the results to changes in those variables where significant uncertainty remains, e.g. price of using ePAQ. Threshold analysis will also be undertaken to assess whether variation in any service characteristics exert significant influence on cost-effectiveness. These results can then be reformulated in a way that is clear to all service providers (Goodacre & Dixon 2005). Sub-group analysis will also be undertaken to assess the cost-effectiveness of using ePAQ in women of different age groups (e.g. under and over 50s) and with clinical conditions (e.g. incontinence and prolapse).

6. Analysis of patient satisfaction

The CSQ-8 is a eight-item self-complete questionnaire, which generates a score with a 8 to 32 range, with a higher score indicating more patient satisfaction. The CSQ-8 is to be posted to patients in both arms of the study, immediately after their fist consultation. A 2 independent samples t-test will be used to compare mean CSQ-8 scores between the groups (ePAQ group vs and usual care) in this parameter. A 95% confidence interval (CI) for the mean difference in this parameter between the ePAQ and usual care groups will also be calculated.

Reducing bias

In this study, the elimination of bias is difficult, as neither patients nor clinicians (AFG & SCR) can be blinded to the group allocations following randomisation. Both AGF and SCR are involved in developing the urogynaecology service and regularly use the ePAQ system as part of routine clinical care. This will be fully acknowledged in any publications or presentations arising from this RCT. Between-group comparisons (AFG & SCR) will be made to look for systematic differences between these 2 consultant clinicians involved in this study. Both control and intervention groups will be given the opportunity to use the ePAQ: this instrument is currently established in routine care in the Urogynaecology Unit and it would be unethical and would potentially introduce bias to withdraw ePAQ use from the control group. Members of the research team (not involved in direct patient care) will carry out recruitment and randomisation of patients as well as the administration of outcome measures. Outcome will be measured in terms of both efficiency and quality, aiming to show whether any changes in efficiency (cost) are related to or at the expense of quality (patient experience) and vice versa. Control patients will use the ePAQ in clinic (as at present) immediately prior to their consultation. The intervention group will complete the questionnaire in advance of a telephone

consultation held with their consultant (AGF or SCR). The use of standardised validated questionnaires as outcome measures (Patient Experience Questionnaire, Client Satisfaction Questionnaire and QQ-10) will also provide generalisable benchmark data for the two groups of patients. Follow-up questionnaires (PEQ, PEI and QQ-10) will be posted to both groups immediately following completion of their first clinical consultation. The micro-costing study will take place over a 2-week period during the second half of the study. Clinicians conducting clinical interviews (in both arms of the study) by telephone and face-to-face will not be involved in issuing, collecting or collating outcome data.

References

1. Jha, S Radley, S,C., Jones, G.L. Patient experience of an electronic questionnaire in clinical practice (abstract). Poster: International Urogynaecology Association AGM, 2007
2. Radley, S,C., Jones, G.L., Tanguy, E.A., Stevens V.G., Nelson, C., Mathers N.J (2006) Computer interviewing in urogynaecology: concept, development and psychometric testing of an electronic pelvic floor assessment questionnaire (ePAQ) in primary and secondary care. British Journal of Obstetrics and Gynaecology 113(2):231-8.
3. *Warner Report, NHS Information Authority, 6th January 2005: Modernisation & New Technology.*
4. Radley, S.C. Jha, S. Jones, G.L. Bates, M. The Virtual Urogynaecology Clinic (abstract). Poster: International Urogynaecology Association AGM, 2007
5. Velikova G, Brown JM, Smith AB, Selby PJ. Computer-based quality of life questionnaires may contribute to doctor-patient interactions in oncology. *Br J Cancer* 2002; 86(1): 51-59.
6. Kleinman L, Leidy NK, Crawley J, Bonomi A, Schoenfeld P. A comparative trial of paper-and-pencil versus computer administration of the Quality of Life in Reflux and Dyspepsia (QOLRAD) questionnaire. Med Care 2001;39(2):181-189.
7. Thornberry J, Bhaskar B, Krulewitch CJ, Wesley B, Hubbard ML, Das A, Foudin L, Adamson M. Audio computerized self-report interview use in prenatal clinics: audio computer-assisted self interview with touch screen to detect alcohol consumption in pregnant women: application of a new technology to an old problem. Comput Inform Nurs 2002;20(2):46-52.
8. Adang RP, Vismans FJ, Ambergen AW, Talmon JL, Hasman A, Flendrig JA. Evaluation of computerised questionnaires designed for patients referred for gastrointestinal endoscopy. Int J Biomed Comput 1991;29(1): 31-44.
9. Arslanian C, Bond M. Computer assisted outcomes research in orthopedics: total joint replacement. J Med Sys 1999;23(3): 239-247.
10. Buxton J, White M, Osoba D. Patients' experiences using a computerized program with a touch-sensitive video monitor for the assessment of health-related quality of life. Qual Life Res 1998;7(6):513-519.
11. Jackson SL, Weber AM, Hull TL, Mitchinson AR, Walters MD. Fecal incontinence in women with urinary incontinence and pelvic organ prolapse. Obstet Gynecol 1997; 89(3):423-427.
12. Snooks SJ, Swash M, Henry M, Setchell M. Risk factors in childbirth causing damage to the pelvic floor innervation. Int J Colorectal Dis 2001;1(1): 20-24.

13. Muth S, Ramsay IN, Pringle S, Marahaj S, Tierney J. The incidence of unvoiced bowel problems in association with urinary incontinence in women. Abstract, ICS (UK) 10th Annual Meeting 2003: 35.
14. NICE Guideline on Incontinence. DoH 2006
15. DoH 2000 Advice on good practice for continence services
16. Curtis L, Netten A. Unit Costs of Health and Social Care 2004. Canterbury: University of Kent, 2004. Department of Health, London; NHS R&D Strategic Review: primary care. Report of Topic Working Group, June 1999.
17. Department of Health, London; Good practice in continence services. NHS
18. Executive, April 2000.Department of Health, London; The NHS Plan: a plan for investment, a plan for reform. NHS Command Paper, July 2000.
19. Department of Health, London; Delivering 21st century IT support for the NHS: national strategic programme. NHS, June 2002.Goodacre S, Dixon S. Is a chest pain observation unit likely to be cost-effective at my hospital? Extrapolation of data from a randomised controlled trial. Emergency Medical Journal 2005;22:418-422.
20. Howie JGR, Heaney DJ, Maxwell M. Measuring quality in general practice. London: Royal College of General Practitioners, 1997 (Occasional paper 75.)
21. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: Developmemt of a general scale. Evaluation and Program Planning 1979; 2: 197-207.
22. Lofland JH, Schaffer M, Goldfarb N. Evaluating health-related quality of life: cost comparison of computerized touch-screen technology and traditional paper systems. Pharmacotherapy 2000;20(11):1390-1395.
23. Moher D, Schulz KF, Altman DG, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomised trials. Lancet 2001; 357: 1191-94.
24. Steine S, Arnstein F, Laerum E. A new, brief questionnaire (PEQ) developed in primary health care for measuring patients' experience of interaction, emotion and consultation outcome. Fam Pract 2001;18(4): 410-18.
25. Steine S, Arnstein F, Laerum E. A new, brief questionnaire (PEQ) developed in primary health care for measuring patients' experience of interaction, emotion and consultation outcome. Fam Pract 2001;18(4): 410-18.
26. M, Brown J, Selby PJ. Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. J Clin Oncol 1999;17(3):998-1007.

Flow Chart for ePAQ study

Letters of referral are received and reviewed by the hospital consultant: SC Radley (SCR) or AG Farkas (AGF)

Eligible patients are telephoned to discuss the study with the research nurse

Interested patients will be sent an information sheet and consent form about the study (appx III)

On receipt of signed consent forms, patients are randomised to either
(1) epaq + telephone consultation (intervention group)
or
(2) usual care (control group)

Allocation will be through stratified block randomisation. SW will generate a randomisation schedule for these strata, using STATA software. This will be held remotely by GJ (who will not be directly involved in the recruitment of patients).

The randomised allocation will then be passed on to the research nurse who will then write to the patient to confirm study entry & outcome of randomisation and arrange appointments

Intervention Group: epaq + telephone consultation

All women in this arm of the study will be sent an information leaflet relating to using epaq-online (appx IV). Women will be asked to complete epaq-online and then asked to telephone to arrange a telephone consultation.

Women who are unable to use the online system will be provided with an appointment to attend hospital for supervised ePAQ completion.

A telephone consultation will take place with the clinician (AGF or SCR), using the questionnaire report, original letter

Control group: usual care

All women in this arm of the study will be given a date and time to attend the urogynaecology clinic

Women will attend the clinic as usual and receive their usual care. This will include being given the opportunity to complete the ePAQ (as is currently normal practice)

On completion of the consultation, all patients will be posted follow up questionnaires (Patient experience questionnaire, Patient enablement Index & QQ-10) and asked to return these in the SAE provided. (appx VI)

6-months following recruitment to the study, all patients will be posted questionnaires relating to their condition and medical care over the preceding 6 months (Appx VII). An SAE will be provided. Non-responders will be re-sent these questionnaires 2 weeks later

Proforma for expression of interest to enter ePAQ study

Research nurse telephones patient and introduces themselves *'My name is ……'* *'I am a research nurse working in Sheffield Teaching Hospitals'*		
'You have been referred to……clinic		
'Would it be possible for me to discuss with you a research study looking at a new form of assessment for women who have been referred to this clinic?'	Yes	No
If yes:		
Confirm the following		
Name …….		
DoB ……...		
Address ……		
Daytime telephone number ……		
Mobile telephone number …….		
Which number would you prefer to be called on?		
'The study is randomised, with 2 groups of women. Half of all women will be first of all asked to complete a questionnaire and their first consultation will then take place shortly afterwards over the telephone with the doctor, before further management is decided. The other half will come to clinic as normal for a routine clinic visit.' *'We will shortly send you some detailed written information in the post about the study.'*		
'Do you have any questions about the study?'		
'Are you possibly interested in helping with this study?'	Yes	No

ePAQ Patient Information sheet

Patient Information Sheet

An invitation to take part in medical research

We would like to invite you to take part in a study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information and, if you wish, discuss it with friends, relatives or your GP. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. If you would like further information please contact the research nurse Hilary Wood on 0114 2268320 or Mr Stephen Radley on 0114 2268167.

What is the purpose of the study?

The ePAQ (electronic Personal Assessment Questionnaire) has been created to assess symptoms and their effect on quality of life in women with pelvic floor problems. It is a computerized system that looks into different aspects of the pelvic floor (including bladder, bowel and vaginal conditions). ePAQ is used routinely in the Urogynaecology clinic to which you have been referred. We have now developed ePAQ further to allow patients to use it before they actually attend for their clinic appointment, in order to allow rapid assessment, early treatment and appropriate clinical follow-up. The purpose of this study is to find out how using ePAQ in this way may affect patient care. This will include looking at the overall time it takes for patients to be assessed, how long it takes for patients to start treatment, what treatments are used and what patients' experiences and feelings are when they have used the system.

What will I have to do if I take part?

If you agree to take part in this study, you will be randomly allocated to one of two groups. Either (1) Using ePAQ or (2) Standard care. This means that there is an equal chance (like the toss of a coin) that you could be in either group. If you are allocated to group (1) you will be asked to use the ePAQ *before* your appointment, following which you will be offered a telephone consultation with your specialist. If you are allocated to group (2), you will be able to use the ePAQ *at the time* of your appointment with your specialist.

If you are allocated to group (1) in this study you will be sent information about how to use the ePAQ system on-line (via the Internet). If you are unable to do this, you will be given an appointment to attend the hospital to complete the questionnaire with help and

supervision. A mutually convenient date, time and telephone number will also be agreed with you for a telephone consultation with your specialist. This consultation will take approximately 15 minutes and will involve discussion of your condition and aims to provide you with information as well as discussing or starting treatment. Appropriate follow-up appointments or investigations will also be arranged for you.

If you are allocated to group (2) in this study, you will be given a routine appointment to attend the Urogynaecology Clinic as normal. You will be given the opportunity to use the ePAQ system at the time of your appointment (as is currently our standard practice in the outpatient department).

We will ask all women who agree to enter this study to complete paper questionnaires about the care that they have received, their quality of life, and their feelings about the service and the ePAQ system that they have used. These paper questionnaires will be posted to you before and immediately following your appointment and at 6 months after your first consultation. They will take about 5 - 10 minutes to complete and you will be provided with a stamped addressed envelope to send them back to us. We will also closely monitor what medical and surgical treatments are used for all women in this study, as well as what investigations are used, what follow-up appointments are arranged and how long patients have to wait.

Why have I been chosen?

Women referred to the Urogynaecology Unit at the Jessop Wing of the Royal Hallamshire Hospital in Sheffield will be given the option of entering this study. We aim to recruit around 304 women to ensure that we can make statistical comparisons between women who use ePAQ in advance of their clinic appointments & then have a telephone consultation, compared with women who use the ePAQ at the time of their appointment.

With your permission, we will inform your GP that you are taking part in this study.

Confidentiality

All reasonable steps will be taken to ensure confidentiality. Any information collected using the Internet ePAQ system is anonymous and can only be used with your permission by the clinicians involved in your care. All the information that is collected about you during the course of the research will be kept strictly confidential. When the results are published, no names will be used, and it will not be possible to identify anyone who has taken part.

Will my taking part involve any extra visits?

If you are allocated to ePAQ (group 1) in this study, you will be asked to complete the ePAQ before your appointment. If you do not have access to the Internet or cannot do this, this will involve an extra visit to the hospital specifically to use the ePAQ. You will be reimbursed for any reasonable travel expenses relating to this visit. You may be able to avoid this visit altogether if you are able to complete the questionnaire via the Internet. The ePAQ takes about 20 – 30 minutes to complete. Once you have completed it, a telephone consultation will be arranged with your specialist. If you undergo a telephone consultation, your specialist may then arrange for you to have tests or receive treatment in a different clinic (other than the Urogynaecology clinic to which you were originally referred). This may mean that you are then seen by a different specialist, physiotherapist or specialist nurse for follow-up. It is also possible that you will be referred back to your GP to start treatment or discharged altogether if this is felt appropriate and you agree. Taking part in this study will not routinely involve any additional or extra tests or investigations. We will however be monitoring patients in this study to see what investigations and treatments are used, to see if there is any effect of using the ePAQ system and telephone consultations on this. There is the possibility that if you are in the ePAQ + telephone consultation group, that you may have your first consultation a little sooner.

What are the possible risks of taking part?

We don't anticipate any health risks as a result of you taking part in this study.

Do I have to take part?

No. Taking part is entirely voluntary. If you prefer not to take part, you do not have to give a reason. Your clinical care will not be affected in any way. If you agree to take part, but later change your mind, you may withdraw at any time without affecting your care in any way.

What will happen to the results of the research study?

The results of this research will help guide future management of patients in the Urogynaecology Department of Sheffield Teaching Hospitals NHS trust and several other NHS trusts throughout the UK. The results will be written up and published in medical journals. We guarantee that it will not be possible to identify you from any report arising from this study. If you would like a copy of the research report we will send this to you.

Who is organizing and funding the research?

Mr Stephen Radley (Consultant Gynaecologist) is organizing the research. The research is funded by a charitable grant from the Sheffield Hospitals Charitable Trust.

What if I wish to complain about the way in which this study has been conducted?

If you have any cause to complain about any aspect of the way in which you have been approached or treated during the course of this study, the normal National Health Service complaints mechanisms are available to you and are not compromised in any way because you have taken part in a research study.

Patient Consent Form

Assessing the impact of using e-PAQ on quality and cost of patient care in urogynaecology: An RCT (STH14733)

PLEASE READ THE FOLLOWING:

IF YOU ARE HAPPY TO TAKE PART IN THIS RESEARCH STUDY PLEASE WOULD YOU COMPLETE YOUR NAME AND ADDRESS AT THE BOTTOM OF THE PAGE AND RETURN THE SIGNED FORM IN THE PRE-PAID ENVELOPE PROVIDED.

PLEASE WRITE **'YES'** OR **'NO'** IN EACH OF THE BOXES BELOW: THANK YOU.

1. I have read and understood the patient information sheet and am happy to participate in this study ☐

2. I understand that I am under no obligation to take part in this research ☐

3. I understand that I am free to withdraw from the study at any time, without giving reason, and without my future care being affected ☐

4. I understand that all the information that is collected will be kept strictly confidential, but that confidentiality may be broken if the investigators feel that there is a risk of harm to myself or others ☐

5. I understand that my GP will be informed of my participation in this research. ☐

Name ……

Date of Birth ……

Address ……..……...

…….

Demographics Proforma

Date completed						
Name					**Postcode**	
Address			**Patient's Date of Birth**			
Ethnic Origin						
White British	**Indian**		**Bangladeshi**			
Black Caribbean	**Chinese**		**Other***			
White Irish	**Mixed Race***		**Pakistani**			
Black African	**White other***		**Black other***		***Please specify**	
Nationality						
Parity						
Educational level	**None**	**Junior**	**Senior**	**University**	**Don't know**	
Marital status	**Married**	**Cohabiting**	**Single**			
Height (M)						
Weight (kg)						

ePAQ Study Telephone Consultation Record

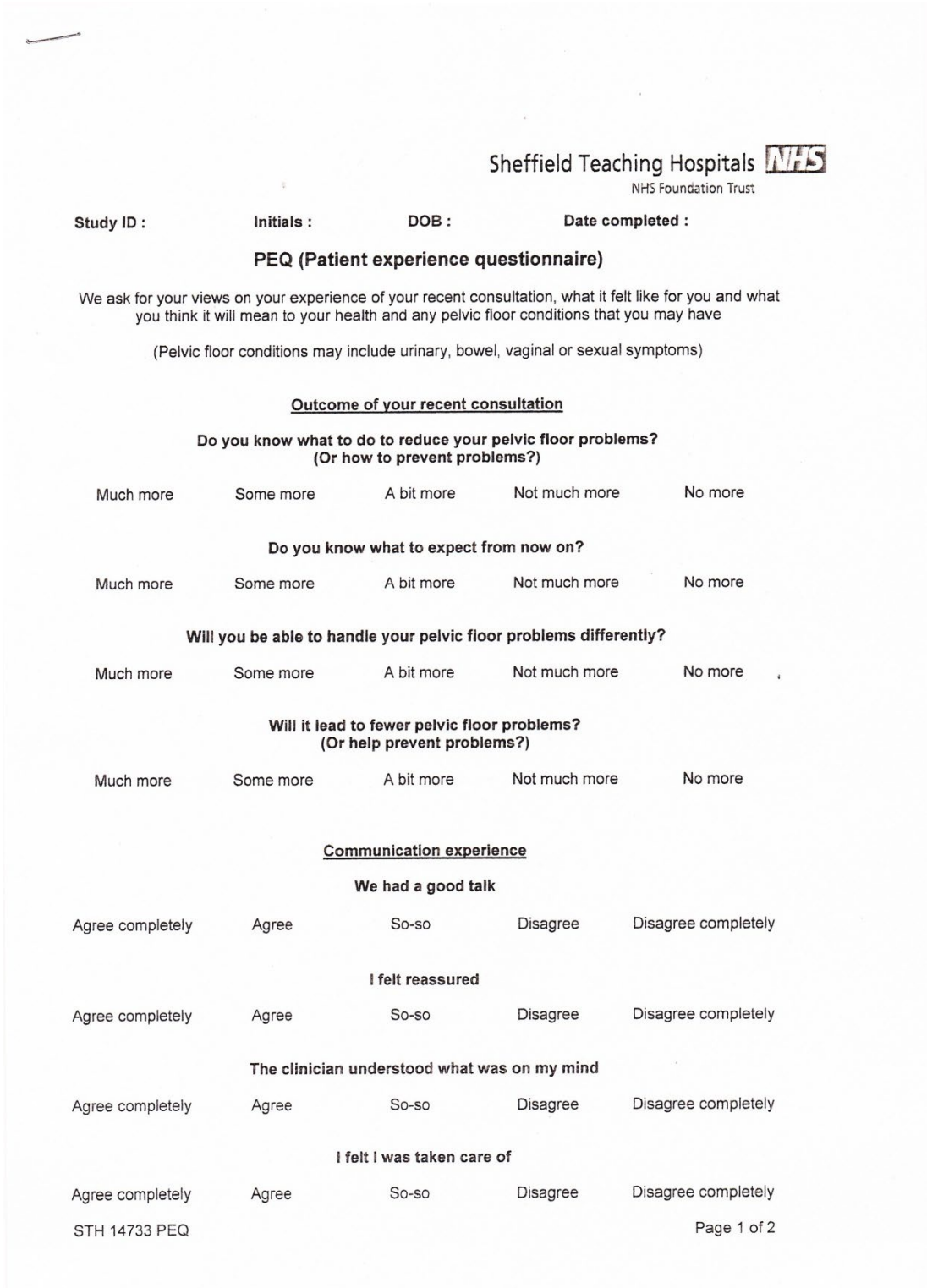
Surname			Unit Number		
First name			DoB		
Address			NHS Number		
e-PAQ Completed	Yes	No	Where ePAQ completed	online	clinic
e-PAQ Downloaded	Yes	No		other:	
e-PAQ Printed	Yes	No	Username		
Date of Telephone Consultation			Questionnaire number		
Telephone Number Dialled			Start Time		
Telephone ext used			End time		
Confirm Name	Yes		No		
Confirm DoB	Yes		No		

Discussion	
Most problematic symptoms	Previous investigations Previous treatment Current treatment
Most prevalent symptoms	Previous investigations Previous treatment Current treatment

Agreed management plan	Investigations				Treatments	
Agreed follow up	Referrals				Follow up appointments	
Information leaflets to be sent	OAB		PFE	SUI	TVT	
	Bladder training		Botox	Urodynamics	Colposuspension	
	Prolapse & You		Pessary	Hysterectomy	Epaq-online	
	Prolapse surgery		Mirena	Contraception	Anaesthesia	
Letter to GP	Yes	No	Letter to Patient	Yes	No	Follow-up:
Letter to Other Clinician	Yes	No	Referral	Yes	No	To:

Appendix 3: ePAQ EEACT Outcome measures

Patient Experience Questionnaire (PEQ)



Sheffield Teaching Hospitals NHS
NHS Foundation Trust

Study ID : **Initials :** **DOB :** **Date completed :**

PEQ (Patient experience questionnaire)

We ask for your views on your experience of your recent consultation, what it felt like for you and what you think it will mean to your health and any pelvic floor conditions that you may have

(Pelvic floor conditions may include urinary, bowel, vaginal or sexual symptoms)

Outcome of your recent consultation

Do you know what to do to reduce your pelvic floor problems? (Or how to prevent problems?)

Much more | Some more | A bit more | Not much more | No more

Do you know what to expect from now on?

Much more | Some more | A bit more | Not much more | No more

Will you be able to handle your pelvic floor problems differently?

Much more | Some more | A bit more | Not much more | No more

Will it lead to fewer pelvic floor problems? (Or help prevent problems?)

Much more | Some more | A bit more | Not much more | No more

Communication experience

We had a good talk

Agree completely | Agree | So-so | Disagree | Disagree completely

I felt reassured

Agree completely | Agree | So-so | Disagree | Disagree completely

The clinician understood what was on my mind

Agree completely | Agree | So-so | Disagree | Disagree completely

I felt I was taken care of

Agree completely | Agree | So-so | Disagree | Disagree completely

STH 14733 PEQ Page 1 of 2

Study ID : Initials :

Communication barriers

It was a bit difficult to connect with the clinician

Agree completely | Agree | So-so | Disagree | Disagree completely

Too much time was spent on small talk

Agree completely | Agree | So-so | Disagree | Disagree completely

It was a bit difficult to ask questions

Agree completely | Agree | So-so | Disagree | Disagree completely

Important decisions were made over my head

Agree completely | Agree | So-so | Disagree | Disagree completely

Emotions immediately after the consultation

After this consultation I felt
(Please circle one number for each line)

Relieved	7	6	5	4	3	2	1	**Worried**
Sad	1	2	3	4	5	6	7	**Cheerful**
Strengthened	7	6	5	4	3	2	1	**Worn out**
Relaxed	7	6	5	4	3	2	1	**Tense**

Client Satisfaction Questionnaire (CSQ-8)

PROTOCOL-STH 145733 COPY - DO NOT USE

CSQ-8 UK English

CLIENT SATISFACTION QUESTIONNAIRE

CSQ-8

Please help us improve our service by answering some questions about the help that you have received. We are interested in your honest opinions, whether they are positive or negative. *Please answer all of the questions.* We also welcome your comments and suggestions. Thank you very much. We appreciate your help.

CIRCLE YOUR ANSWERS

1. How would you rate the quality of service you received?

4 *Excellent*	3 *Good*	2 *Fair*	1 *Poor*

2. Did you get the kind of service you wanted?

1 *No, definitely not*	2 *No, not really*	3 *Yes, generally*	4 *Yes, definitely*

3. To what extent has our service met your needs?

4 *Almost all of my needs have been met*	3 *Most of my needs have been met*	2 *Only a few of my needs have been met*	1 *None of my needs have been met*

4. If a friend were in need of similar help, would you recommend our service to him or her?

1 *No, definitely not*	2 *No, I don't think so*	3 *Yes, I think so*	4 *Yes, definitely*

5. How satisfied are you with the amount of help you received?

1 *Quite dissatisfied*	2 *Indifferent or mildly dissatisfied*	3 *Mostly satisfied*	4 *Very satisfied*

6. Have the services you received helped you to deal more effectively with your problems?

4 *Yes, they helped a great deal*	3 *Yes, they helped somewhat*	2 *No, they really didn't help*	1 *No, they seemed to make things worse*

7. In an overall, general sense, how satisfied are you with the service you received?

4 *Very satisfied*	3 *Mostly satisfied*	2 *Indifferent or mildly dissatisfied*	1 *Quite dissatisfied*

8. If you were to seek help again, would you come back to our service?

1 *No, definitely not*	2 *No, I don't think so*	3 *Yes, I think so*	4 *Yes, definitely*

WRITE ANY COMMENTS OVERLEAF

TMS.002

QQ-10 Questionnaire

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

Study ID: **Initials:** **DOB:** **Date completed:**

QQ-10

Please circle the answers below each of the following 10 statements that best fit your feelings about the computer questionnaire (ePAQ) that you recently completed

Please use the boxes at the bottom of the next page to make additional comments.

The questionnaire helped me to communicate about my condition

Strongly agree | Mostly agree | Neither agree or disagree | Mostly disagree | Strongly disagree

The questionnaire was relevant to my condition

Strongly agree | Mostly agree | Neither agree or disagree | Mostly disagree | Strongly disagree

The questionnaire was easy to complete

Strongly agree | Mostly agree | Neither agree or disagree | Mostly disagree | Strongly disagree

The questionnaire included all the aspects of my condition that I am concerned about

Strongly agree | Mostly agree | Neither agree or disagree | Mostly disagree | Strongly disagree

I enjoyed filling in the questionnaire

Strongly agree | Mostly agree | Neither agree or disagree | Mostly disagree | Strongly disagree

I would be happy to complete the questionnaire again in the future as part of my routine care

Strongly agree | Mostly agree | Neither agree or disagree | Mostly disagree | Strongly disagree

Study ID: **Initials:**

The questionnaire was too long

Strongly agree	Mostly agree	Neither agree or disagree	Mostly disagree	Strongly disagree

The questionnaire was too embarrassing

Strongly agree	Mostly agree	Neither agree or disagree	Mostly disagree	Strongly disagree

The questionnaire was too complicated

Strongly agree	Mostly agree	Neither agree or disagree	Mostly disagree	Strongly disagree

The questionnaire upset me

Strongly agree	Mostly agree	Neither agree or disagree	Mostly disagree	Strongly disagree

Do you have any additional comments or suggestions to make about the ePAQ questionnaire?

Were any of your important symptoms, problems or concerns missed out?

Do you have any comments or suggestions about the service as a whole?

SF-12

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

Study ID: Initials: DOB:

The SF-12v2™ Health Survey

Your Health in General

Please answer every question. Some questions may look like others, but each one is different. Please take time to read and answer each question carefully by ticking what best represents your answer.

1. In general, would you say your health is:

Excellent	Very good	Good	Fair	Poor

2. The following questions are about activities you might do during a typical day. Does **your health now limit you** in these activities? If so, how much?

	Yes, Limited a lot	Yes, limited a little	No, not limited at all
a) **Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf			
b) Climbing **several** flights of stairs			

3. During the **past 4 weeks**, how much of the time have you had any of the following problems with your work or other regular daily activities **as a result of your physical health**?

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
a) **Accomplished less** than you would like					
b) Were limited in the **kind** of work or other activities					

STH 14733

Study ID: Initials: DOB:

4. During the **past 4 weeks**, how much of the time have you had any of the following problems with your work or other regular daily activities **as a result of any emotional problems** (such as feeling depressed or anxious)?

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
a) **Accomplished less** than you would like					
b) Did work or other activities **less carefully than usual**					

5. During the **past 4 weeks**, how much did **pain** interfere with your normal work (including both work outside the home and housework)?

Not at all	A little bit	Moderately	Quite a bit	Extremely

6. These questions are about how you feel and how things have been with you during the **past 4 weeks**. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the **past 4 weeks**...

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
a) have you felt calm and peaceful?					
b) did you have a lot of energy?					
c) have you felt downhearted and blue?					

7. During the **past 4 weeks**, how much of the time has your **physical health or emotional problems** interfered with your social activities (like visiting friends, relatives, etc.)?

All of the time	Most of the time	Some of the time	A little of the time	None of the time

THANK YOU TAKING COMPLETING THIS QUESTIONNAIRE.

STH 14733

Appendix 4: ePAQ EEACT Ethics Committee Approvals

NHS

National Research Ethics Service

CM/SR

13 November 2007

Mr Stephen Radley
Consultant in Obstetrics & Gynaecology
Sheffield Teaching Hospitals NHS FT
Jessop Wing
Royal Hallamshire Hospital
Sheffield
S10 1NT

15 NOV 2007

Dear Mr Radley

Full title of study: **Assessing the impact of using ePAQ on quality and cost of patient care in urogynaecology: an RCT (STH14733)**

REC reference number: **07/H1308/130**

Thank you for your email dated 30 October 2007, responding to the Committee's request for further information on the above research.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation [as revised].

Ethical review of research sites

The Committee has designated this study as exempt from site-specific assessment (SSA. There is no requirement for [other] Local Research Ethics Committees to be informed or for site-specific assessment to be carried out at each site.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Application		12 September 2007
Investigator CV		
Protocol	1	28 May 2007

Covering Letter	1	19 October 2007
Summary/Synopsis	1	28 May 2007
Questionnaire: QQ-10	1	28 May 2007
Questionnaire: PEI	1	28 May 2007
Questionnaire: PEQ	1	28 May 2007
Questionnaire: SF012		
Participant Information Sheet	2	19 October 2007
Participant Consent Form	3	30 October 2007
Response to Request for Further Information	2	30 October 2007
Response to Request for Further Information	1	19 October 2007
Email correspondence	1	30 October 2007
ePAQ telephone consultation record	1	28 May 2007
Demographics proforma	1	28 May 2007
Proforma for expression of interest	1	28 May 2007

R&D approval

All researchers and research collaborators who will be participating in the research at NHS sites should apply for R&D approval from the relevant care organisation, if they have not yet done so. R&D approval is required, whether or not the study is exempt from SSA. You should advise researchers and local collaborators accordingly.

Guidance on applying for R&D approval is available from http://www.rdforum.nhs.uk/rdform.htm.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Website > After Review

Here you will find links to the following

a) Providing feedback. You are invited to give your view of the service that you have received from the National Research Ethics Service on the application procedure. If you wish to make your views known please use the feedback form available on the website.
b) Progress Reports. Please refer to the attached Standard conditions of approval by Research Ethics Committees.
c) Safety Reports. Please refer to the attached Standard conditions of approval by Research Ethics Committees.
d) Amendments. Please refer to the attached Standard conditions of approval by Research Ethics Committees.
e) End of Study/Project. Please refer to the attached Standard conditions of approval by Research Ethics Committees.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nationalres.org.uk .

07/H1308/130	Please quote this number on all correspondence

With the Committee's best wishes for the success of this project

Yours sincerely

[redacted]

Dr C A Moore
Chair

Email: sue.rose@sth.nhs.uk

Enclosures: *Standard approval conditions SL-AC2 for other studies*

Copy to: Professor Simon Heller
[R&D office for NHS care organisation at lead site]

NHS

National Research Ethics Service

North Sheffield Local Research Ethics Committee
1st Floor Vickers Corridor
Northern General Hospital
Herries Road
Sheffield
S5 7AU

Tel: 0114 2714894
Fax: 0114 256 2469

04 September 2008

Mr Stephen Radley
Consultant in Obstetrics & Gynaecology
Jessop Wing
Royal Hallamshire Hospital
Sheffield
S10 1NT

- 8 SEP 2008

Dear Mr Radley

Study title:	**Assessing the impact of using ePAQ on quality and cost of patient care in urogynaecology: an RCT (STH14733)**
REC reference:	**07/H1308/130**
Amendment number:	**1**
Amendment date:	**07 August 2008**

The above amendment was reviewed at the meeting of the Sub-Committee of the REC held on 27 August 2008.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Questionnaire	SF12v2 Health Survey	11 August 2008
Protocol	3	07 August 2008
Participant Information Sheet	3	07 August 2008
GP Letter	2	07 August 2008
Notice of Substantial Amendment (non-CTIMPs)	1	07 August 2008
Covering Letter	1	07 August 2008

Membership of the Committee

The members of the Committee who were present at the meeting are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

07/H1308/130:	Please quote this number on all correspondence

Yours sincerely

[redacted]

Ms Sue Rose
Committee Co-ordinator

E-mail: sue.rose@sth.nhs.uk

Enclosures — List of names and professions of members who were present at the meeting and those who submitted written comments

Copy to: *Sheffield Teaching Hospitals NHS Trust*

North Sheffield Local Research Ethics Committee
Attendance at Sub-Committee of the REC meeting on 27 August 2008

Dr C M Moore Chair	Consultant anaesthetist
Miss P Kingman (Lay member)	Retired
Mrs S Rose	Coordinator

Appendix 5: ePAQ EEACT Clinical Outcome Results

This appendix presents the clinical study results from the ePAQ study. The clinical analysis was led by Dr Georgina Jones, and completed by Richard Jacques.

Table 34. **Unadjusted + adjusted differences in mean Patient Experience Questionnaire scores at 6 months by treatment groups**

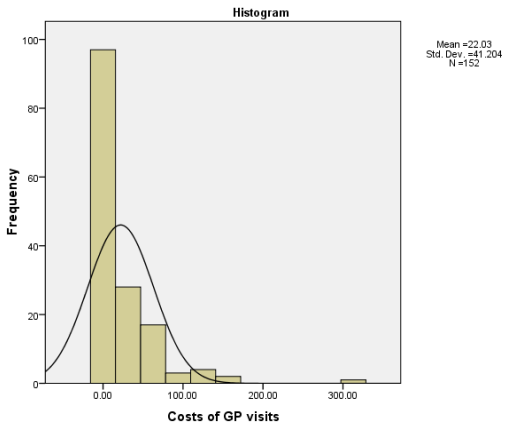
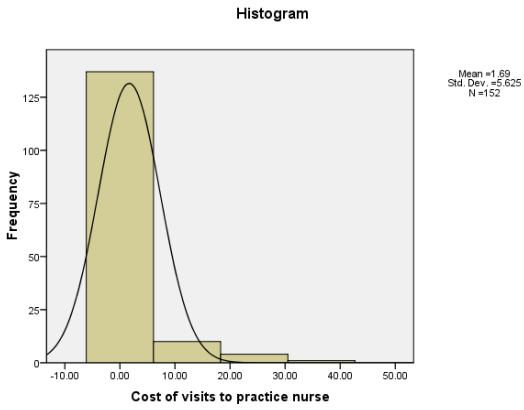
	Control			Intervention			Unadjusted[2]				Adjusted[3]			
								95% CI				95% CI		
PEQ dimension[1]	N	Mean	SD	N	Mean	SD	Mean Diff	Lower	Upper	P-Value	Mean Diff	Lower	Upper	P-Value
Short-term outcome	69	2.9	1.1	68	2.9	1.2	-0.04	-0.4	0.3	0.843	-0.03	-0.4	0.4	0.887

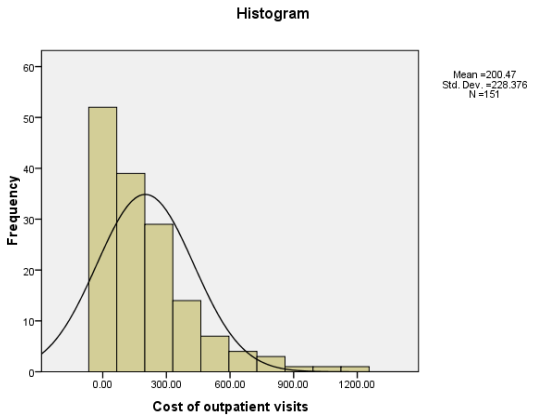
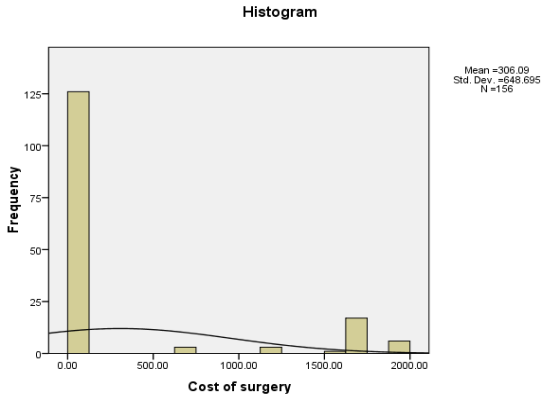
	Control			Intervention			Unadjusted[2]				Adjusted[3]			
								95% CI				95% CI		
PEQ dimension[1]	N	Mean	SD	N	Mean	SD	Mean Diff	Lower	Upper	P-Value	Mean Diff	Lower	Upper	P-Value
Communications	68	3.8	0.8	68	4.2	0.6	0.4	0.2	0.6	0.001	0.4	0.2	0.7	0.001
Emotions	66	4.6	1.3	64	5.4	1.2	0.7	0.2	1.1	0.001	0.7	0.3	1.1	0.002
Barriers	68	4.0	0.6	67	4.4	0.6	0.3	0.1	0.5	0.002	0.3	0.1	0.5	0.003

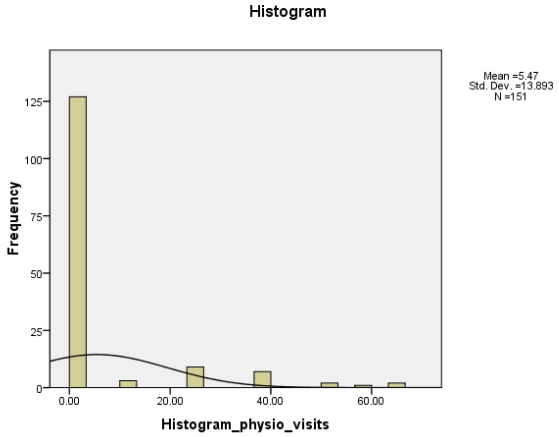
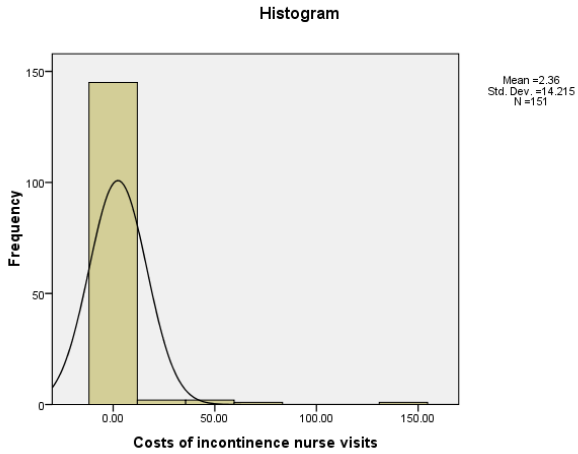
[1]The PEQ dimensions are scored on a 1-5 scale with the exception of emotions which is 1-7. A high score represents a good communication experience, positive emotions, positive consultation outcome and a lack of communication barriers. [2]P-value from independent samples t-test. [3]Adjusted mean difference calculated from a linear regression model with PEQ dimension score as the outcome and age, parity and treatment group as covariates. A positive mean difference indicates that the intervention group has the better score.

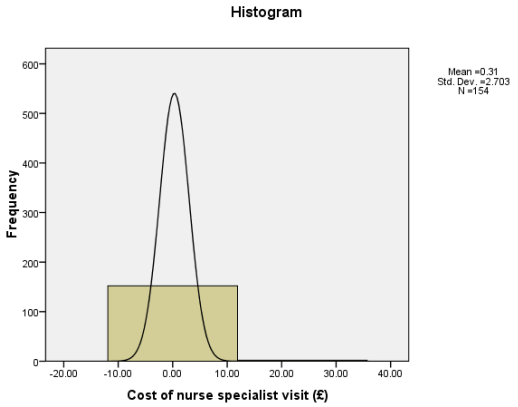
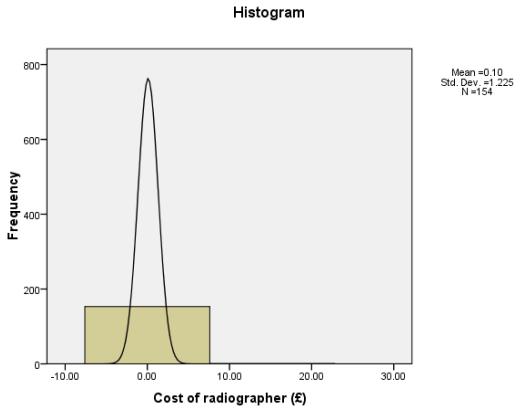
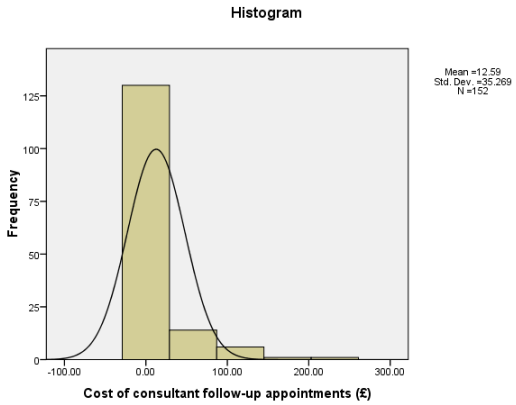
Appendix 6: Measures of distribution of ePAQ trial cost data

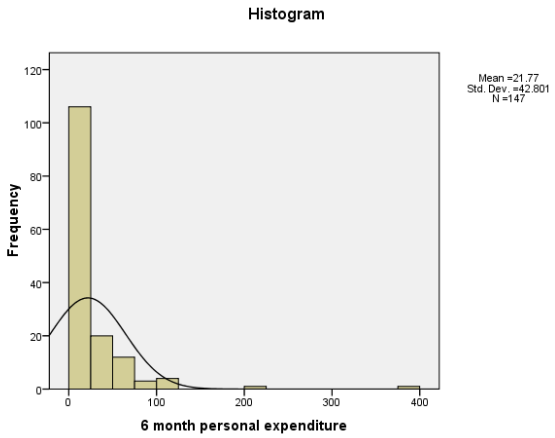
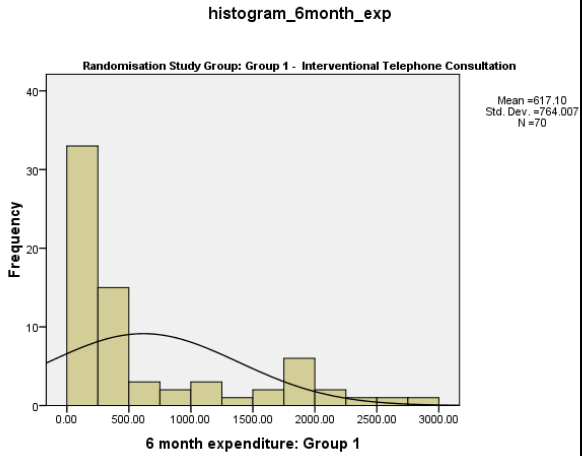
Table 35. Measures of distribution of ePAQ trial cost data.

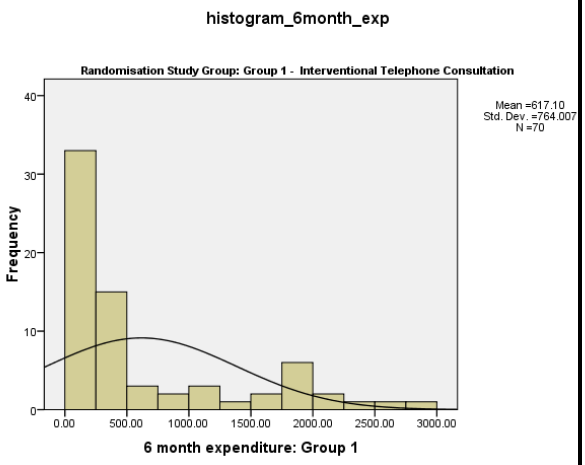
Cost Item	Histogram	Summary
GP Visits		Non-normally distributed Skewness= 3.369 Kurtosis = 17.018
Nurse Visits		Non-normally distributed Skewness= 3.78 Kurtosis = 15.35

Cost Item	Histogram	Summary
Outpatient visits		Non-normally distributed Skewness= 1.66 Kurtosis = 3.3
Surgical Procedures		Non-normally distributed Skewness= 1.743 Kurtosis = 1.187

Cost Item	Histogram	Summary
Physio-therapy visits		Non-normally distributed Skewness= 2.615 Kurtosis = 6.15
Incontinence nurse		Non-normally distributed Skewness= 7.766 Kurtosis = 68.39

Cost Item	Histogram	Summary
Nurse specialist		Non-normally distributed Skewness= 8.69 Kurtosis = 74.45
Radiographer		Non-normally distributed Skewness= 12.41 Kurtosis = 154.00
Consultant re-visit		Non-normally distributed Skewness= 3.44 Kurtosis = 13.62

Cost Item	Histogram	Summary
6 month personal expenditure		Non-normally distributed Skewness= 5.44 Kurtosis = 41.49
Total 6 month resource use Intervention group		Non-normally distributed Skewness= 1.321 Kurtosis = .462

Cost Item	Histogram	Summary
Total 6 month resource use Control group		Non-normally distributed Skewness= 1.525 Kurtosis = .865

Appendix 7: ePAQ EEACT Poster Presentation

RTI(*h*)(*s*)

An Economic Evaluation Alongside a Clinical Trial in Pelvic Floor Medicine

Victoria K Brennan,[1] Simon Dixon,[2] Georgina L Jones,[2] Stephen Radley,[3] Richard Jacques,[2] HJ Wood,[2] William Ledger[4]

[1] RTI Health Solutions, Sheffield, United Kingdom; [2] ScHARR, University of Sheffield, United Kingdom; [3] Sheffield Teaching Hospitals Charitable Trust, United Kingdom; [4] University of New South Wales, Sydney, Australia

BACKGROUND

- ePAQ-PF is a web-based interactive questionnaire (electronic pelvic floor assessment questionnaire) that provides a detailed evaluation of a woman's pelvic floor symptoms and their impact on her quality of life. The information is used by the consultant as an adjuvant to their consultation.
- An online version of ePAQ-PF allows women to access and complete the questionnaire prior to their clinic appointment, enabling the information to be used to triage patients.
- ePAQ-PF has facilitated development of telephone consultations in place of clinic visits. The consultant can gain a clear and concise summary of a patient's signs and symptoms and the impact on her life, without having to undertake a prolonged consultation that includes questions some women might find embarrassing to answer.

OBJECTIVE

- To determine the cost-effectiveness of using an online questionnaire (ePAQ-PF) in combination with a telephone consultation compared with standard care.

METHODS

Patient Recruitment

- All women referred to Sheffield Teaching Hospitals urogynaecology services, aged > 18 years, able to read and understand English were eligible.
- From June 2008 until December 2009, potential participants were identified by a consultant review of referral letters received from general practitioners (GPs) or other clinicians.
- Eligible patients were contacted by telephone to discuss the study. Patients who expressed an interest in participating were sent a study information leaflet and consent form.

Randomisation

- On receipt of a signed consent form, patients were randomised to one of two groups:
  - **Control:** Women were mailed an appointment at the clinic, with the option of completing the ePAQ-PF on arrival, immediately before the face-to-face consultation.
  - **Intervention:** Women were posted an information letter and leaflet and a voucher inviting them to complete ePAQ-PF online. Results supported a subsequent telephone consultation.

Resource Use

- Resource use data were collected at two study time points:
  - Initial consultation: including staff time, computer and software, and overhead.
  - Six months following study recruitment: Resource use data collected via a postal questionnaire included personal expenditure due to bladder, bowel, or vaginal problems; time taken off work; time away from usual activities; visits to health professionals; and inpatient and outpatient visits.

Unit Costs

- United Kingdom (UK) unit costs were applied to resource use estimates for patients in each study arm. All costs were inflated to 2011 prices using the pay and prices index.[1]
- Staff time was costed using the unit costs of health and social care, including overhead costs.[1] Costs of surgical procedures were estimated using HRG data.[2]
- A micro-costing study was undertaken on a subset of patients to derive costs for the initial ePAQ-PF completion and consultation. This included all patients receiving their consultation (telephone or face-to-face) during one of three set months throughout the trial. The duration of these consultations was timed. For patients attending a face-to-face consultation, computer costs were calculated assuming two touch-screen computers with an average of 1,000 completions each per year with replacement every 5 years. Annual software costs were calculated assuming an average of 1,000 completions per year. These data were provided by the lead clinic consultant.

Utilities

- The SF-12 was administered to patients at baseline and 6 months after initial consultation.
- Results from the SF-12 were converted into utilities using the SF-6D.[3]

Cost-effectiveness

- Estimates of mean costs and quality-adjusted life-years (QALYs) for the two treatment arms were calculated over the 6-month follow-up period.
- The study was taken from a societal perspective. Due to a 6-month time horizon, total costs and QALYs were not discounted.
- Analysis of the SF-6D results was performed only on patients who had completed the SF-12 at both baseline and 6 months.
- Cost-effectiveness was determined using the incremental cost-effectiveness ratio (incremental costs divided by incremental benefit).

Patient Experience

- Patient experience was measured by the Patient Experience Questionnaire (PEQ) and Client Satisfaction Questionnaire (CSQ).

RESULTS

- 195 women were randomised; 98 to the intervention group (92 received the intervention); 97 to the control group (95 completed the study). See Table 1.
- Table 2 reports resource use estimates for each group. The independent t-test was used to compare the mean values from the two study arms.[4]

Table 1. Baseline Characteristics of Women Randomised to Each Trial Arm

		Control (n = 97)		Intervention (n = 98)	
		n	Mean (Range)	n	Mean (Range)
Age (years)		95	50.6 (24.7–77.1)	97	51.5 (28.8–75.1)
Height (m)		93	1.63 (1.47–1.90)	97	1.63 (1.50–1.78)
Weight (kg)		93	70.4 (45.5–122.0)	97	71.9 (48.0–112.0)
Body mass index		92	26.6 (19.4–46.1)	97	27.2 (19.7–45.8)
Parity		97	2 (0–5)	91	2 (0–6)
Nationality	American	0	(0%)	1	(1%)
	British	95	(98%)	88	(91%)
	European	1	(1%)	1	(1%)
	Irish	1	(1%)	1	(1%)
Ethnic origin	White British	95	(98%)	86	(95%)
	White Irish	1	(1%)	1	(1%)
	White other	0	(0%)	3	(2%)
	Mixed race	1	(1%)	1	(1%)
Education level	None	17	(17%)	14	(15%)
	Junior	26	(27%)	20	(22%)
	Senior	27	(28%)	21	(23%)
	University	27	(28%)	25	(28%)
Marital status	Married	71	(73%)	64	(70%)
	Cohabiting	8	(8%)	8	(9%)
	Single	2	(2%)	4	(4%)
	Widowed	5	(5%)	3	(3%)
	Divorced	8	(8%)	10	(11%)
	Separated	3	(3%)	1	(1%)
	Civil partnership	0	(0%)	1	(1%)

Table 2. Resource Use

	Intervention		Control		Mean Difference (95% CI)	P Value
Resource	n	Mean (SD)	n	Mean (SD)		
GP visits	72	0.119 (0.112)	78	0.050 (0.102)	-0.113 (-0.272 to 0.320)	0.500
Practice nurse	72	0.169 (.524)	79	0.190 0(0.577)	-0.002 (-0.169 to 0.164)	0.977
Outpatient visits	74	1.590 (1.937)	77	1.440 (1.520)	0.153 (-0.404 to 0.710)	0.589
Mean number of surgical procedures	74	0.220 (0.414)	78	0.190 (0.398)	0.037 (-0.092 to 0.165)	0.572
Other professionals						
Physiotherapist	74	0.400 (1.039)	77	0.450 (1.142)	-0.056 (-0.406 to 0.294)	0.753
Stoma nurse	74	0.013 (.115)	79	0.000 (0.000)	0.074 (-0.013 to 0.940)	0.321
Incontinence nurse	74	0.135 (.775)	77	0.000 (0.254)	0.044 (-0.149 to 0.236)	0.655
Nurse specialist gynaecology	76	0.039 (.161)	78	0.000 (0.000)	0.026 (-0.011 to 0.063)	0.159
Consultant (GP)	75	0.119 (.601)	77	0.260 (0.016)	-0.066 (-0.261 to 0.129)	0.362
Societal						
Personal expenditure follow-up period (£)	74	27.944 (53.207)	72	15.503 (21.666)	12.441 (-1.355 to 25.294)	0.079
Time off work (days)	76	3.090 (12.472)	79	5.210 (20.752)	-2.119 (-7.703 to 3.525)	0.459
Time away from normal activities	74	2.410 (8.029)	79	1.290 (5.752)	1.114 (-1.252 to 3.480)	0.354

- Estimated mean costs per patient are reported in Table 3. Consultation costs for the intervention group were less than half those of the control group, driven primarily by the duration of the consultation and associated labour costs.
- The intervention group incurred greater direct costs and personal expenditure during follow-up. However, lower costs associated with productivity loss for the intervention group resulted in lower indirect costs per patient.
- Mean total costs per patient were £30.04 greater in the intervention group than the control group.
- SF-6D reduced slightly during follow-up for the intervention group and increased slightly for the controls, resulting in QALY loss for the intervention group and QALY gains for the controls (Table 4).
- PEQ results suggest a statistically significant difference between the treatment groups for the communication, emotions, and barriers domains, although this was not significant for the short-term dimension (Table 4). Patient satisfaction as measured using the CSQ showed no significant differences between the control and intervention groups (not shown).
- Figure 1 shows the cost-effectiveness acceptability curve, which indicates that when the value of the ceiling ratio is between £20,000 and £30,000, the probability of cost-effectiveness is between 30% and 40%. The shape of the curve reflects the intervention being slightly more costly, less effective, and with a large amount of uncertainty around these estimates.

Table 3. Mean Costs Per Patient (£)

Resource	Intervention (n = 92)	Control (n = 95)
Cost of consultations		
Consultation	29.35	69.52
Software	2.40	2.40
Computer	N/A	0.25
Total per patient	31.75	72.17
Direct costs during 6-month follow-up		
GP visits	41.22 (45.40)	35.33 (65.79)
Practice nurse	0.94 (3.40)	2.13 (5.08)
Outpatient visits	250.07 (316.26)	186.30 (246.547)
Surgical procedures	330.44 (737.275)	265.63 (594.36)
Other professionals		
Physiotherapist	5.04 (12.59)	4.35 (15.75)
Specialist nurse	4.52 (18.25)	2.32 (11.14)
Consultant (in person)	77.54 (24.44)	14.25 (27.30)
Total direct costs	646.77 (844.46)	522.41 (662.05)
Indirect costs during 6-month follow-up		
Personal expenditure	24.07 (31.05)	16.17 (20.37)
Loss of productivity	443.26 (1,572.15)	461.07 (1,475.01)
Total indirect costs	467.33 (1,565.42)	457.24 (1,475.25)
Total costs per patient	1,129.26 (2,102.24)	1,101.82 (2,120.44)

Table 4. Mean Utility and PEQ Estimates per Patient

	Intervention		Control		Mean Difference (SE) [95% CI]	P Value
Item	n	Mean (SD)	n	Mean (SD)		
SF-6D baseline	27	0.64 (0.90)	30	0.62 (0.091)	0.026 (0.024) (-0.022 to 0.07441)	0.287
SF-6D 6 months	27	0.63 (0.092)	30	0.62 (0.091)	0.00999 (0.023) (-0.035 to 0.05214)	0.703
Change in SF-6D	27	-0.0152 (0.072)	30	0.0039 (0.094)	-0.01999 (0.02245) (-0.06257 to 0.02500)	0.401
QALYs gained	27	-0.0076 (0.037)	30	0.0019 (0.047)	-0.0095 (0.11122) (-0.31199 to 0.012300)	0.401
PEQ: Short-term outcome	69	2.5 (0.1)	68	2.5 (0.2)	-0.04 (-0.4 to 0.3)	0.843
PEQ: Communication	68	3.8 (0.8)	68	4.2 (0.6)	-0.04 (-0.4 to 0.3)	0.843
PEQ: Emotions	66	4.6 (0.3)	64	5.4 (0.2)	0.4 (0.2 to 0.6)	0.001
PEQ: Barriers	68	4.0 (0.6)	67	4.4 (0.6)	0.7 (0.2 to 1.1)	0.001

Figure 1. Cost-effectiveness Acceptability Curve

DISCUSSION

Although the intervention was not cost-effective compared with the controls, there was a significant difference in an important aspect of the care process—the communications, emotions and barriers domains of the PEQ—which was not captured by the ICER. This highlights the importance of decision makers accounting for intervention effects that fall outside the conventional conceptualisation of the QALY. Methods could be developed that allow effects not directly related to health, such as process utility, to be incorporated into the QALY.

CONCLUSIONS

Statistically significant gains in patient experience were identified for the intervention group, although in strict cost-utility terms, the intervention was dominated by the control. Incremental costs and QALYs resulted in a negative incremental cost-effectiveness ratio.

REFERENCES

Please see handout for complete reference list.

DISCLOSURE

Sheffield Teaching Hospitals Charitable Trust provided funding for this project.

CONTACT INFORMATION

Victoria Brennan, MSc, RTI Health Solutions
Phone: +44 (0) 1142 122291; E-mail: vbrennan@rti.org

Presented at: ISPOR 16th Annual European Congress
November 2-6, 2013
Dublin, Ireland

Appendix 8: Process Utility Manuscript Copyright confirmation.

Copy of email requesting confirmation of copyright permissions for inclusion of manuscript in thesis, and response:

Dear Victoria,

I think this is the relevant text:

http://www.springer.com/gp/open-access/authors-rights/self-archiving-policy/2124

So my interpretation is that you can include the final accepted versions of the papers in your thesis (i.e. the final revised version you submitted to the journal).

BW, Chris

-----Original Message-----

From: SpringerAlerts@springeronline.com

[mailto:SpringerAlerts@springeronline.com]

Sent: Friday, 11 September 2015 10:44 p.m.

To: Carswell, Chris, Springer

Subject: Author question from springer.com

Dear Chris,

I had 2 papers published in Pharmacoeconomics, which were part of my PhD studies. I am now writing my thesis and would like to include them in the appendix. However, I am not sure if the copyright rules allow it. I require a letter from the Journal to include in my thesis giving me permission.

The papers are:

Brennan VK and Dixon S. (2013) Incorporating process utility into quality adjusted life years: a systematic review of empirical studies. *Pharmacoeconomics*. Aug; 31 (8),677-91.

Brennan VK and Dixon S. (2014) Response to letter to editor: Capturing disutility from waiting time. *Pharmacoeconomics*. Apr; 32(4),421-2.

Please can you advise?

Many thanks

Victoria

Appendix 9: Process Utility Manuscript Submission Covering Letter

RTI Health Solutions
Velocity House
3 Solly Street
Sheffield S1 4DE
UK

Editor In Chief

Pharmaco-economics

10th May 2013

Dear Chris Carswell,

RE: PECA-D-13-00049R2, entitled "**Incorporating process utility into quality adjusted life years: a systematic review of empirical studies**."

Many thanks for the further review of our manuscript and for allowing us to re-submit following amendments. We have addressed the comments, and this includes an update of the literature searches.

Please find attached an updated manuscript (tracked and clean version), an updated PRISMA diagram (tracked and clean version), and a document summarising the comments and actions.

If you require any additional information on the manuscript please do let me know.

Best Wishes,

Victoria Brennan

Senior Health Economist, RTI-HS UK,

Email: vbrennan@rti.org

Appendix 10: Process Utility Manuscript

The final publication is: Brennan, V. and Dixon, S. (2013) Incorporating process utility into quality adjusted life years: a systematic review of empirical studies. Pharmacoeconomics. Aug;31(8):677-91, and is available at Springer via http://dx.doi.org/ 10.1007/s40273-013-0066-1.

Incorporating process utility into quality adjusted life years: a systematic review of empirical studies.

Victoria K Brennan[1§], MSc; Professor Simon Dixon, PhD[2].

[1§] RTI-Health Solutions, Velocity House, Business and Conference Centre, 3 Solly Street, Sheffield S1 4DE UK. Phone: +44 (0) 1142 133391. Fax: +44 (0)161 434 8232.

[2] HEDS, ScHARR The University of Sheffield Regent Court, 30 Regent Street Sheffield S1 4DA UK

§ Corresponding author

Email Addresses:

VKB: vbrennan@rti.org

SD: s.dixon@shef.ac.uk

Word count: 5,956 (exclusive of title, abstract, references and figure legends).

Acknowledgements: Victoria K Brennan and Simon Dixon received no funding for the preparation of this article and have no conflicts of interest that are directly relevant to the content of this review.

ABSTRACT

Objective: This review aimed to identify published studies which provide an empirical measure of process utility which can be incorporated into estimates of QALY calculations.

Methods: A literature search was conducted in PubMed to identify published studies of process utility. Articles were included if they were written in the English language and reported empirical measures of process utility which could be incorporated into the QALY calculation; those studies which reported utilities which were not anchored on a scale of 0 representing dead and 1 representing full health were excluded from the review.

Results: Fifteen studies published between 1996 and 2012 were included. Studies included respondents from the United States of America, Australia, Scotland and the United Kingdom (UK), Europe and Canada. Eight of the included studies explored process utility associated with treatments; six explored process utility associated with screening procedures or tests; and one was performed in preventative care.

A variety of approaches were used to detect and measure process utility: 4 studies used standard-gamble techniques; 4 studies used TTO; 1 study used conjoint analysis and 1 used a combination of conjoint analysis and TTO; 1 study used SF-36 data; 1 study used both TTO and EQ-5D; and 3 studies used wait-trade-off techniques.

Measures of process utility for different drug delivery methods ranged from 0.02 to 0.27. Utility estimates associated with different dosing strategies

ranged from 0.005 to 0.09. Estimates for convenience (able to take on an empty stomach) ranged from 0.001 to 0.028. Estimates of process utility associated with screening and testing procedures ranged from 0.0005 to 0.031. Both of these estimates were obtained for management approaches to cervical cancer screening.

Conclusion: The identification of studies through conventional methods was difficult due to the lack of consistent indexing and terminology across studies; however, the evidence does support the existence of process utility in treatment, screening and preventative care settings. There was considerable variation between estimates. The range of methodological approaches used to identify and measure process utility, coupled with the need for further research into, for example, the application of estimates in economic models means it is difficult to know whether these differences are a true reflection of the amount of process utility which enters into an individual's utility function, or whether they are associated with features of the studies methodological design. Without further work, and a standardised approach to the methodology for the detection and measurement of process utility, comparisons between estimates are difficult. This literature review supports the existence of process utility and indicates that, despite the need for further research in the area, it could be an important component of an individual's utility function which should at least be considered, if not incorporated into cost-utility analyses.

Key points for decision makers

- Reported empirical studies consistently support the presence of process utility within the provision of health care.
- There is a lot of heterogeneity in the study methods and results which prevents any robust conclusions to be made about the estimates of process utility
- Without a direct comparison of methods within a single study, with clearer tests of validity, it is unclear which are the best methods for estimating process utility
- A failure to incorporate process utility into the QALY calculation may mean that economists are reporting cost-utility analyses which present inaccurate estimates of cost-effectiveness
- Identification of best practice methods would benefit future researchers and decision makers.

Introduction

Within health economics, it is common practice to seek to maximise health gain within a given budget constraint. Health therefore needs to be measured and valued, and in economic evaluations, this is frequently achieved through the estimation of quality-adjusted-life years (QALY). [1] In England and Wales, the use of QALY's is now required by the National Institute for Health and Clinical Excellence (NICE) and the All Wales Medicines Strategy Group (AWMSG) as an outcome for health technology assessments and is used as a common outcome measure to enable comparisons across different areas of health care. Use of the QALY remains one of the key outcome measures

for the proposed value based pricing (VBP) system in the UK.[2] In order to calculate QALYs utilities are required. These are a measure of preference for specific states of health quantified on a cardinal scale anchored by 0 representing “dead” and 1 representing “full health”. Utility estimates are combined with the number of years spent in the health state to generate QALYs.[3]

The reliance on the QALY within the UK is founded on the notion of extra-welfarism that has been forwarded by prominent health economists as being relevant to the decision making context of a publically funded health service.[4] However, if a welfarist framework, or less restrictive brand of extra-welfarism were adopted, then valuations may not always be restricted to health outcomes alone, but could also encompass factors such as the effect of the process of receiving care or “process utility” on patients. These are not routinely incorporated into UK analyses as it is argued that the health consequences experienced by the patient are the most important element of the patient’s health care utility function.[5] This is reflected in the multi-attribute preference based measures (PBM) commonly used to estimate utilities (e.g. the EQ-5D, SF-6D), which attempt to derive a function for the estimation of utility of different health states, depending on different attributes of health (e.g. mobility, self-care, usual activities, pain/discomfort, anxiety/depression). Although we may not be surprised that the primary focus of health service evaluation is on health outcomes, there is a body of literature questioning whether this introduces narrow consequentialism (defined by Mooney as “ the monopolization of the utility of welfare function....by health and health alone”[6]) and has led to researchers ignoring

other consequence and process issues whose inclusion would take into account that the process of receiving care impacts on patients' utility, as well as the health gains they achieve.[5] A recent research paper examining the operationalisation of VBP also suggests that, although health improvement is the most important benefit of medicines, there are process-related aspects of healthcare, such as "being treated with dignity,……being treated in a way that the patient finds less unpleasant, e.g. oral as compared with intravenous administration…" which also matter to people.[2]

The literature provides no consensus on what elements should be included in the utility function to avoid narrow consequentialism, and so inherently provide a more holistic evaluation of a consumer's utility. However, several suggestions on the components to include in the utility function comprise an element of "process"; for example, McAlister[7] suggested the inclusion of three dimensions: the process of care; the health outcome; and the structure of the providers of care (for example, including the resources, physical and organisational settings), and Donaldson and Shackley [8] suggest including processes (e.g. non-health aspects of care, such as receiving surgery or drugs), other health attributes (e.g. the effects of health care on life years and functional health related quality of life), and health attributes such as dignity and re-assurance.

There is also disagreement on defining what aspects of care fall into the "process utility" category; Donaldson & Shackley [8] define non-health aspects of care such as receiving surgery or drugs as process utility. Swan et al [9] imply a wider definition by suggesting that "testing and other short-

term events can also be viewed in terms of “process utility” which attaches utility or value to aspects that are proximal to eventual outcomes, such as information, reassurance, and morbidity.” A broader term of non-health outcomes has also been used, and formed the basis of a review by Opmeer and colleagues[10]. Their review identified process related preferences as just one type of non-health outcome, with intervention related outcomes (e.g. invasiveness) and patient related outcome (e.g. uncertainty) being seen as separate types. Whichever approach is adopted to define the pleuristic benefits of health care (encapsulating both health and process attributes) a failure to embrace this more holistic approach to utility measurement allows health economists to place restrictions on what enters consumers’ utility functions. [9]

Even wider utility functions that incorporate notions of equity have also been explored[11,12,13] Whilst such approaches move beyond individual values to social welfare functions, they do highlight the fact that the specific ‘brand’ of economic evaluation that has developed within health technology assessment represents not only a single type of evaluation (i.e. cost-utility analysis) but also incorporates a narrow consequentialist approach to utility valuation. As such, we recognise that other economic evaluation frameworks exist that can capture broader sets of outcomes, for example, cost-benefit analysis or cost-consequences analysis.[14] Within these frameworks, and indeed cost utility analysis, other methods of benefit valuation are possible that seek to capture patient preferences. Opmeer and colleagues identified nine different preference elicitation methods used in the literature to value non-health outcomes; binary choice, judgment of separate

components, trade-off, discrete choice experiments, likert scale, visual analogue scale, accept in future situation, ranking, and recommendation to others[10]. From an economic evaluation perspective, it is also informative to divide studies into just two categories; those which provide methods of identifying and quantifying process utility which can be incorporated into the QALY calculation, and can therefore be used within cost-utility analyses (CUA), and those which cannot. One of the most commonly used approaches for CUA incompatible studies is "willingness to pay" (WTP) where, for example, participants may be presented with two treatment options, and are then asked how much they would be willing to pay for the more favourable alternative. For example, Lloyd et al [15] found that patients with diabetes were willing to pay approximately £35-£45 for a reduction from 3 injections per day to 1.[16] Whilst such studies provide a measure of the monetary value patients place on the dose frequency attribute, these values cannot be incorporated into QALY estimates (without further analysis).

This dichotomisation of studies is important as many economic evaluation guidelines around the world recommend the use of cost-utility analysis in health care. If evaluations are to meet these guidelines, then process utility must be captured within QALY estimates. So, whilst a focus on studies that generate values that can be incorporated into a QALY may appear narrow, it directly addresses a specific and important global policy issue. Consequently, the following review aims to identify published studies which provide an empirical measure of process utility which can be incorporated into the QALY calculation and therefore be used as input parameters for

[16] Values of WTP read from Figure within Lloyd et al., (2011), exact values are not reported (13).

CUA. The review explores the methodological approaches taken to incorporate process utility into the utility function. This is considered to be an essential first step in identifying a research agenda to develop best practice.

Methods

Literature Search

We searched PubMed, an electronic database of biomedical literature for papers which reported an empirical measure of process utility. Searches were performed in April 2013. The search strategy was based on 4 combinations of free text terms; these are reported in Table I. We initially used a search strategy based on the term “process-utility” (search strategy 1). However, as we knew several published studies did not use this term, particularly in drug and treatment administration, we applied 3 additional combinations of free-text terms to the search: Search strategy 2 combined “treatment related attributes” with utility or “utility measurement”; Search strategy 3 combined “dose frequency” with utility or “utility measurement”; and the final search strategy combined “drug administration” with utility or “utility measurement”. In addition, the reference list of identified articles and reviews were checked for studies meeting the inclusion criteria. The final list of identified studies was also circulated to experts in the field requesting the identification of any additional studies they were aware of not identified by the searches.

Inclusion Criteria

Any study, published in the English language, which reported an empirical measure of process utility which was anchored on a scale of 0 representing dead and 1 representing full health, which could therefore be incorporated into the QALY was included in the review. No further exclusion or date restrictions were applied.

The study selection process was performed by one researcher who initially screened the titles and abstracts for eligibility (VB). The full text versions of all included studies were then obtained and reviewed by eligibility by one researcher using the same inclusion and exclusion criteria. Any uncertainty about the inclusion of studies was checked by a second researcher (SD).

Data collection and extraction and synthesis.

Data were extracted by one researcher into standardised data extraction tables (Tables II-IV) which contained items about general study characteristics, details on the measurement of process utility and process utility estimates. General study characteristics included the year of publication and country of study, the clinical indication (for example, cervical screening, haematology treatment or preventative health care), and the intervention (for example, inhaled versus injectable treatments, breast biopsy or antenatal screening). Methods of the measurement of process utility were characterised by a description of the source of process utility (for example, flexible dosing strategies, route of administration or open versus laparoscopic surgery), a description of the hypothetical health states where

appropriate, study population (for example, patients or general population), and details on the valuation methodology, including the anchoring of death where used. Separate tables were used to report process utility gains from drug delivery methods/dosing strategies (Table III), and process utility gains from screening/texting procedures (Table IV). Any uncertainty surrounding the extraction of data was checked with second researcher.

Results

The literature search results are presented in Figure 1. Following removal of duplicates, the PubMed and hand searches identified 99 studies which were screened for eligibility. An additional 1 was identified through the author's (SD) knowledge of the field. A total of 15 studies met the inclusion criteria and were included in the review.

Comparison of key characteristics.

A summary of the key study characteristics can be seen in Table I.

Year of publication and country of analysis

The studies identified and reviewed were published during the years 1994 to 2012. Five studies were performed in the United States of America (USA), 4 in Australia, 2 in Scotland, 1 in the United Kingdom (UK), and 2 were multi-country studies, one of which included respondents from Europe and the USA and the other included respondents from Canada, Australia and the USA.

Indication.

Eight of the included studies were performed in specific treatments; 3 of these were related to diabetes specific treatments [21, 22, 23,], one explored process utility within surgical treatments for gallstone disease [16], one explored the use of opioid treatment in chronic pain 17], one was performed in HIV related treatments [20], one in the treatment of schizophrenia [19], and one in haematology. [18] A further six studies explored the presence and measurement of process utility in the screening/testing setting; two of these were performed in cervical screening [24, 25]; of the remaining 4 studies, one was completed in antenatal screening [26]; one study was performed in breast cancer biopsy [28]; one explored process utility in cerebrovascular testing [9], and one in testing in patients with peripheral vascular disease [27]. The remaining study was performed in preventative care [29].

Population

Ten of the included studies used patients to estimate health state utilities. The remaining 5 studies were performed using the general population. No studies provided estimates from health care professionals.

Methodological characteristics

A variety of approaches were used to detect and measure process utility: 4 studies used standard-gamble techniques [23, 24- 26]; 4 studies used TTO [16, 18, 19, 21, 29]; 1 study used conjoint analysis [17] and 1 used a combination of conjoint analysis and TTO [22]; 1 study used SF-36 data [20]; 1 study used both TTO and EQ-5D valuation of hypothetical health states [21]; 3 studies used waiting-trade-off techniques. [9, 27, 28]

Anchoring of death

Several of the included studies used a two-stage or chaining technique for the valuation of health states, specifically when the states under evaluation do not contain a risk of death. [16, 21, 23, 24, 25] Instead of immediate death being compared to full health, an "anchor" state which is better than death but worse than the scenarios under evaluation is used. This health state is then compared with an intermediate health state of interest. Scores are then converted onto a scale of 0 representing death and 1 representing full health. This is the methodology suggested by Torrance [30] for the valuation of temporary health states.

Overview of study results

All studies identified the presence of process utility. Table III shows the results of studies evaluating process utility associated with treatments. Measures of process utility for different drug delivery methods ranged from 0.02 [21] to 0.27 [18]. Utility estimates associated with different dosing strategies ranged from 0.005 [21] to 0.09 [19]. Estimates for convenience (able to take on an empty stomach) ranged from 0.001 to 0.028 [20].

Estimates of process utility associated with screening and testing procedures can be seen in Table IV. Measures ranged from 0.0005 [25] to 0.031 [24]. Both of these estimates were obtained for different management approaches to cervical cancer screening.

Descriptions of individual studies

The earliest of the included studies was performed by Cook et al [16] who undertook a CUA comparing extracorporeal shock wave lithotripsy (ESWL) to open and laparoscopic cholecystectomy for gallstone disease. ESWL is a

non-invasive day-case procedure with minor levels of post-procedure morbidity, but diarrhoea and nausea are experienced by some patients. A laparoscopic cholysectomy is a minor operation, and patients can usually return to normal activities 2 to 3 days post-operatively. There is a small probability of serious complications; however, post-operative morbidity is significantly reduced when compared to open surgery. Respondents were asked to value hypothetical health states using a two stage SG technique. The author's postulate that the short duration of the disutility experienced post surgery may be overshadowed by the disutility arising from the process of receiving surgery and therefore having the associated risk of death. To overcome this they asked respondents to quantify disutility for the operation (ex ante) and the riskless health state (ex post). The results were summed to provide a "*partial ex ante*" perspective. In essence the *ex ante* evaluation of the operation represents the "process" of receiving care, and the *ex post* evaluation of the health states refers to the health outcomes.[16]

The authors calculated a total loss of QALYs associated with the three procedures from both an ex post perspective, and an ex ante perspective. The results indicated that the ex ante approach lead to considerably higher QALY loss when compared to the ex post approach. The cost per QALY indicated that the open cholecystectomy was always dominated, but that the choice of approach (*ex ante* vs *ex post*) had a significant impact on the cost-effectiveness of the laparoscopic cholectsyectomy and ESWL. Where an *ex post* approach was taken laparoscopic cholecystectomy dominated. When an *ex ante* approach was used the QALY loss from laparoscopic cholecystectomy was greater than the loss associated with ESWL. [16] This

study clearly indicates that the perspective taken and items included in the utility function can have a significant impact on estimates of cost-effectiveness.

Schmier et al. [17] explored the presence of process utility in the treatment of chronic pain in patients receiving opioid analgesia. They used conjoint analysis techniques to explore patient preferences for four key treatment attributes: pain control; side effects; severity of side effects; and the route of administration (oral versus patch). The latter of which refers to a “process” of health care. The study included 96 patients with non-malignant pain and 25 patients with malignant pain. Those patients with malignancy were also asked to complete the QLQ-C30 and patients with non-malignant pain were asked to complete the MOS SF-36. A computer assisted method of interviewing was used which automatically converted participants’ responses to utility estimates. The results showed differences in the mean utilities assigned to health states that differed only by oral versus patch route of administration. For patients with non-malignant pain there was a small and consistent preference for the oral administration of opioids. For patients with malignant pain preferences differed by severity of side effects, with patients preferring oral administration when pain was mild and patch administration when pain was severe. Utility differences between the same health states but receiving medication by oral versus patch medication ranged from no difference (for severe respiratory depression health state) to 0.16 in preference for oral medication for the health state describing moderate respiratory depression. [17]

Osborne et al. [18] conducted a time-trade-off (TTO) study to explore the utility associated with subcutaneous infusion when compared to the oral administration of iron chelation therapy. They used TTO techniques to estimate preferences for three hypothetical health states with differing treatment modalities but identical treatment outcomes: an anchor state where the patient has iron chelation but the treatment itself is not described; the anchor state plus treatment via a subcutaneous infusion; and the anchor state plus treatment via once-daily oral treatment. They found a mean difference of 0.23 (median 0.27) in the utility estimates for the two treatment health states in preference for oral administration. This was found to be statistically significant [18].

Osborne et al.[19] used a similar approach to estimate utility for alternative treatment intervals for long acting antipsychotic injections for the treatment of schizophrenia. They used TTO techniques to estimate preferences for 4 hypothetical health states. The first health state described untreated/relapsed schizophrenia; the remaining health states had identical clinical outcomes but differed in the intervals between injections: once every 2 weeks, once every 4 weeks and once every 3 months. The results of the study showed that utility increased as the time between injections increased (2-weekly injections: mean utility = 0.61; 4 weekly injections: mean utility = 0.65; 3 monthly injections: mean utility = 0.70).

Kauf et al. [20] used SF-36 responses from five clinical trials in HIV to explore the marginal impact on utilities of different dosing frequencies in HAART treatment in patients with HIV. Dosing frequencies were categorised as one

or two times per day, with twice per day (BID) as the reference level. Food and drink requirements were also explored (whether medication must be taken with food and/or drink or on an empty stomach). A total of 1327 participants were included in the study. Small differences between utility estimates according to dosing strategies were identified: for medication which had to be taken with food -0.028; medications to be taken on an empty stomach -0.001; medications taken once per day 0.020. [20]

Several studies have explored process utility associated with treatment attributes in patients with diabetes. Chancellor et al. [21] explored patient preferences for inhaled versus injectable insulin regimens. Written descriptions were developed for five clinical scenarios (two for type 1 and three for type 2 diabetes), with two alternative insulin scenarios described for each: injectable-only or inhaled insulin to replace or reduce the number of daily injections. Equal efficacy was assumed within each scenario pair. 344 patients participated in the study, with each rating scenario pairs associated with their diabetes type, and their current health using both TTO techniques and the EQ-5D. A majority of patients preferred inhalation to injectable treatment. Differences in utilities in favour of inhaled administration ranged from 0.04 to 0.09 using TTO techniques, and 0.02 to 0.07 using the EQ-5D. [21]

Polster et al. [22] compared patients' preferences for attributes of two diabetes treatments. These attributes included comparisons of once daily versus twice daily doses. They used a survey instrument consisting of the EQ-5D, a TTO exercise where respondents compared the profiles of the two treatments,

and a conjoint exercise in which respondents compared a series of TTO exercises for hypothetical product profiles. Within the TTO task the authors chose to use “life free of diabetes” in place of the anchor for “perfect health” to more closely reflect a realistic decision scenario for respondents, and to limit the impact of potential co-morbidities. The results showed a 0.005 increase in utility for movement from a twice daily dose frequency to a once a day frequency.[22]

Boye et al. [23] explored the differences in utility associated with three injection-related attributes in patients with diabetes. Two of these were related to process: dose frequency and dose flexibility. They asked patients with type 2 diabetes to value hypothetical health states and their own current health using a two stage SG technique. A total of 151 patients completed interviews. They found that weekly injections were associated with an average added utility of 0.023 when compared to daily injections. Dose flexibility compared medication which can be taken at any time of the day to medication taken with food. They found an added average utility of 0.006 for medication taken at any time of the day. [23]

Studies have also been undertaken to evaluate process utility associated with screening/testing procedures. [24, 25] Birch et al. [24] compared aggressive versus conservative management approaches to the follow-up of mildly abnormal Pap smears. They asked women attending family planning clinics to evaluate 6 hypothetical scenarios which differed in management approach but had the same health outcomes using a two-stage SG technique. They found that for comparisons where no abnormalities were detected, an

immediate colposcopy (normal) versus observation followed by 3 repeat smears compared to an immediate colposcopy resulted in an increase in utility of 0.031. For comparisons where mild abnormalities were found, they compared 3 repeat smears followed by cryotherapy to an immediate colposcopy followed by cryotherapy, and identified an associated -0.021 difference in mean utilities in favour of immediate colposcopy followed by cryotherapy. For comparisons where abnormalities were found which necessitated a cone biopsy, they compared 3 repeat smears followed by biopsy to an immediate colposcopy followed by biopsy and identified a -0.017 difference in utility estimates in favour of immediate colposcopy followed by biopsy.[24]

Howard et al [25] also explored process utility associated with alternative approaches to the testing and management of abnormal atypical squamous cells in women. They asked women to evaluate four clinical management scenarios, using two-stage SG. The scenarios either described the management of a lesion that resolved spontaneously following either a repeat Pap smear or negative HPV test; or scenarios describing the management of a lesion that required treatment following either a repeat Pap smear or a positive Pap test. The results found that there was evidence of process utility, with HPV testing strategies having lower valuations than repeat Pap smear tests where the clinical outcome was the same. These results indicate that including measurements of process utility in studies would have implications on management decisions, as although it is suggested that HPV triage testing can offer a more effective management strategy over conventional management for women with atypical squamous

cells, the women themselves do not necessarily see HPV testing as having higher utility than conventional management with repeat smear.[25]

Cairns et al. [26] used SG techniques to estimate the value placed on improved information in antenatal screening for cystic fibrosis carrier status. Two methods of administering the cystic fibrosis carrier status screening test were evaluated; stepwise screening and couple screening. Stepwise screening was initially offered to the pregnant woman, only if the test is positive was the test offered to her partner. In couple screening both partners were tested at the same time, and the result of the test combined the outcome. If they had a positive test, they are then offered prenatal diagnosis, where the cystic fibrosis status of the foetus was determined. This incurred approximately a 1 in 200 risk of foetus loss. Women were asked to trade off the risk of fetal loss with improved information from diagnostic testing. The women were initially presented with the scenario where there was a 1 in 100 chance that the foetus had cystic fibrosis. For the remaining 5 scenarios the SG methods were used to establish the risk of foetal loss associated with amniocentesis at which the woman was indifferent as to whether or not the diagnostic test was performed. The results of the study indicated as the risk of detecting a cystic fibrosis infected foetus increased, the greater are the potential benefits of a diagnostic test. For both stepwise screening and couple screening, the expected utility of screening was found to be greater than that of no screening. . When considering preferences for either method of screening: at the group level there was no statistically significant difference between the expected utilities of the two methods; at the individual level , stepwise screening was preferred overall, however, couple screening

was preferred when considering only those women who preferred screening to no screening.[26]

Swan et al. [27] used a variation of the TTO methodology to assess patient's preferences for magnetic resonance angiography (MRA) versus x-ray angiography (XRA) in patients with peripheral vascular disease (PVD). They use wait trade off techniques (WTO) by asking the patient to trade off extended time with the condition being diagnosed in order to avoid the adverse effects of one screening test compared to another. They suggest this technique can produce utility estimates which can be used to determine QALYs. They found significant differences between patients' wait times and corresponding disutilities for MRA versus XRA which resulted in a difference of 9.36 QALDs between tests and equates to a 0.026 difference in utility estimates in favour of MRA. [27]

Swan et al. [9] used the same WTO methodology to explore process utility associated with imaging in cerebrovascular disease. They compared the preferences of 89 patients for conventional versus magnetic resonance (MR). Patients had experienced both tests. They calculated a weighted difference between the time a patient was willing to wait for a test result and treatment after a hypothetical "ideal" test and the choice to undergo conventional angiography or MR angiography with immediate treatment. The study found a significant difference between patient preferences for conventional and MR angiography in favour of MR, with a utility increment of 0.014. [9]

Swan et al. [28] again used WTO techniques to investigate process utility associated with different methods of biopsy for the detection of breast cancer. They compared preferences for core-needle biopsy (CNB) with the more invasive excisional surgical biopsy (EXB) in 38 women. The study indicated that women were willing to wait for significantly longer to avoid the EXB. They used patient's baseline and test-related anxiety, measured using TTO, to scale the WTO. Median TTO preferences for baseline was 1.00 and anxiety was 0.93. The QALY toll (the amount of quality adjusted life expectancy one is willing to give up to avoid an experience) for EXB was 1.5 and for CNB was 0.5, this equates to a utility value of 0.001 in favour of CNB. [28]

The final included study explored the presence of process utility in the prevention of hip fractures. Salkeld et al. [29] asked participants to value 6 scenarios using TTO. The scenarios included a "full health" state, two process states (wearing hip protectors and feeling discomfort; wearing hip protectors and feeling confident), and three states associated with psychological outcomes of regret and elation. Utility estimates for the process states were 0.31 for discomfort compared to 0.64 for confidence. [29]

DISCUSSION AND RECOMMENDATIONS FOR FUTURE RESEARCH.

This review was undertaken to identify published studies which provide an empirical measure of process utility which can be incorporated into the QALY calculation and therefore be used as input parameters for cost-utility analyses.

Although 15 studies were identified through the searches, the identification of studies through conventional methods was difficult due to the lack of consistent indexing and terminology across the studies. This resulted in several studies being identified through ad hoc methods, and the authors acknowledge that there may be additional studies which the searches failed to identify. Those studies which were identified do demonstrate that whilst there is no universally accepted methodology for the detection and measurement of process utility, the evidence does support its existence. However, when drawing conclusions from the review it is important to consider the extent of publication bias, where studies resulting in negative findings are less likely to be published or to receive pharmaceutical funding to be pursued to publication.

Estimates

The review identified studies associated with processes involved in the treatment of patients (through both drug administration and surgical procedures), preventative care (prevention of hip fractures) and the screening of individuals for the prevention and early detection of morbidities. The utility estimates showed considerable variation, with estimates for treatment processes ranging from 0.001 for both comparisons of a laparoscopic cholecystectomy with an EWSL procedure in gallstone disease [16] and as a decrement for medications which have to be taken on an empty stomach [20], to 0.09 for both inhaled versus injectable medication in diabetes [20], and for receiving 2 weekly versus 3 monthly injectable medication in schizophrenia[19]. There was also variation between measurements of the same processes identified within different studies. For

example, both Kauf et al. [20] and Boye et al. [23] compared patient preferences for taking medication which needs to be taken with food or an empty stomach and medication which can be taken irrespective of the meal times. Boye et al. [23] estimated a utility increment of 0.006 for medication which can be taken at any time of the day, this compares to Kauf et al's [20] decrements of 0.0028 for medication which must be taken with food, and 0.001 for medication which must be taken on an empty stomach. Due to the range of methodological approaches used to identify and measure process utility it is difficult to know whether these differences in estimates are a true reflection of the amount of process utility which enter into an individual's utility function, or whether they are associated with features of the studies methodological design. Without a standardised approach to the methodology for the detection and measurement of process utility comparisons between estimates are difficult and should be viewed with caution.

The statistical and clinical significance of some of the estimates needs to be considered. Not all differences identified reached conventional levels of statistical significance and are below the minimally important difference in utilities suggested by Walters and Brazier , which ranges from 0.010 to 0.048 (as measured on the SF-6D range)[31]. Conversely, one of the larger estimates does not appear plausible when compared to existing valuations of various conditions on the utility scale (see for example, Stouthard and Colleagues, 2001)[32]. For example, it is hard to believe that patients really consider the difference between subcutaneous and oral drug administration (process utility 0.27[18]) equivalent to be similar in magnitude to the difference between perfect health and for instance anorexia, severe hearing

disorder or permanent cognitive impairment after bacterial meningitis, which have utilities between 0.70 and > 0.80.[13] (We thank one of the referees for highlighting both of these points).

Study Populations

The populations used to value health state utilities varied. Ten of the included studies used patients, and the remaining 5 used a general population sample. Research into the area indicates that utilities assigned to the same health states vary substantially according to the group of individuals valuing them, with patients typically producing higher values than the general population. [33] This added heterogeneity between studies introduces additional uncertainty to the process utility decrements identified.

As well as variability of results, the validity of the study populations also needs to be considered. Washington Panel and NICE methods guidelines for instance, reflect a strong theoretical preference for values to be based on those of the general population.[34] Yet the majority of the studies that we examined, and that were reviewed by Opmeer, used patient respondents[10]. There may be an argument that the specialised nature of the treatment information can only be fully appreciated by patients, however, given the systematic differences seen between patient and general population values[35] , the use of patient values in process utilities makes their results incompatible with Washington Panel and NICE reference case methods.

The studies included were performed in varied therapeutic areas, suggesting that process utility is relevant to a substantial number of patients. However, most were performed in chronic conditions (diabetes [21, 22, 23], chronic pain

[17], HIV [20], schizophrenia[19] and haematology[18]), and screening (cervical screening [24, 25], antenatal screening [26]; breast cancer biopsy [28]; cerebrovascular testing [9] peripheral vascular disease [27]), with the remaining study being performed in preventative care[29]. The focus of research appears to be on conditions where the treatment burden is greatest, and hence, where therapeutic developments have been targeted on more convenient treatments. Whilst the concept of process utility is relevant to all conditions, the need to develop convenient therapeutic solutions to acute treatments is less pressing.

Methods

Several of the included studies [16, 21, , 23, 24, 25] used a two stage or chaining methodology for the valuation of health states, where intermediate health states are measured relative to the best health state and the worst state . This is a methodology put forward by Torrance [30] to overcome issues associated with the valuation of temporary health states, and leads onto an additional point for consideration: the issue of how the process utility estimates are applied within cost-effectiveness analyses. [36] In order to apply the utility estimates for example in an economic model, assumptions must be made on whether the effects are deemed to be persistent (and so extend beyond the actual process), or whether they are transient (and so apply only to the time during which the process occurs). [37] An example of this may be a cost-effectiveness analysis into the use of inhaled insulin versus injectable insulin. Applying the utility gain only during the time which insulin is administered versus its application to the associated health state will have a massive effect on the estimated QALY gains from the improved

process. If the former approach is taken then the QALY gains for some processes may be very small. One of the studies by Osborne [18] for example, identifies a utility decrement of 0.27 associated with moving from subcutaneous to oral drug administration, which seems implausible for an ongoing reduction in utility, but possible for a utility reduction on the day of treatment. Whilst the majority of studies used recognised TTO and SG techniques for the valuation of health states, one study by Kauf et al. [19] used responses from the SF-36, a generic PBM, to determine estimates of process. This raises the point that the extent by which process utility is not captured by PBMs is unknown, and is an important issue which needs further research. If process utility is captured by existing PBM's then adding further allowances to the QALY estimates will lead to a double counting of its effects. It is therefore important that the process insensitivity of preference based measures is ruled out empirically before any separate estimates of process utility are added into cost-effectiveness analyses.

It is also important to consider the presence and impact of focussing effects, whereby respondents valuing health states place too much emphasis on the component of interest – in this case the element of process – which may in turn lead to an exaggeration of the associated utility decrement. [36] This has much in common with the broader criticisms of developing disease specific health state scenarios. One approach to overcome is to bolt-on process measures to generic PBM's. This is a methodology which has been used in the past, however, these studies have been restricted to the addition of further health domains (cognition, sleep and pain) [38-40]. It would be possible to add on bespoke process descriptions or descriptions from a validated

measure of patient experience, for example, the patient experience questionnaire (PEQ).[41]

The use of process utilities within cost-utility analyses

The assumption within this paper is that process utilities can be added into the QALY calculation that underpins a cost-utility analysis in the same way as a health effect. We have already mentioned that this requires the decision maker to accept a broader concept of wellbeing than that typically adopted within extra-welfarism. But beyond this philosophical issue, other technical issues need to be considered.

One thing that is not clear in the studies examined above is the time period used within many of the trade-off questions. This is important as it needs to be consistent with the time period to which the utility is applied within the QALY calculation. Identifying a utility decrement associated with an invasive test that lasts for one hour of 0.5, needs to be applied to one hour (expressed as a fraction of a year) within the cost-utility analysis. Applying short-lasting disutilities to chronic health states is clearly wrong.

One issue that has been raised within the evaluation of more convenient therapies is that they generate gains in process utility and compliance, with the latter then generating further utility gains through greater effectiveness. So, does the inclusion of process utility double-count the benefits of the new therapy? We would argue that it does not; process and therapeutic effects are quite different. However, if as part of the valuation, patients are considering the added therapeutic benefits, then double counting will occur.

As such, future studies need to guard against this in much the same way that income effects related to ill-health are excluded from health state utilities.

Future Research

More rigorous testing of the validity of the methods used is essential in future studies. This should take the form of tests of validity within the stated preference studies themselves. For example, the impact of framing, alternative elicitation methods and the impact of the size/number of process attributes would assess the degree to which utility values suggest possible biases.[42]

Patient based studies should be undertaken to examine the extent to which PBMs do, or do not capture process utilities. Much of the literature takes the insensitivity of PBMs as a given, yet the results of Kauf and colleagues suggest that with sufficiently large sample sizes, PBMs may be capable of detecting small utility differences.[19]

Also within patient based studies, comparisons of process utility estimates with patient reported outcome measures relating to treatment satisfaction or burden should be undertaken to assess the degree of convergent validity with utility values.. Interestingly, the study of inhaled insulin by Chancellor was 'tested' through the launch of the related product – Exubera – and observed patient uptake.[21] Whilst the stated preference study suggested important utility gains from moving to inhaled insulin, patients were very reluctant to use it and the product was eventually withdrawn from the market. In other words, patient's revealed preferences suggested a utility *decrement* associated with using inhaled insulin.

The inhaled insulin estimates may have reflected problems with the research methods used, but could also be due to the poor prediction of experienced utility by participants. Dolan and Kahneman [43] suggest that our ability to adapt to changed circumstances such as ill-health is not accounted for in the valuation of health states by healthy individuals. Whilst this critique is related to all stated preference studies, it may be that the nature of the changes being assessed in process utility studies are particularly prone to this problem. Treatment inconvenience is a very emotive issue in an increasingly consumerist sector, and as such, may have just transient effects on a patient's welfare. The degree to which changes in treatment patterns are initially resisted, then accepted by patients, should be examined. If it is found that anticipated utility behaves in a dramatically different manner for process characteristics, as opposed to health characteristics, then the case for adding them into the QALY would be significantly weakened.

The implications of incorporating process utility into economic analyses on health care policy should be carefully considered. It is evident that non-health outcomes are important to patients, and therefore should not be ignored in the setting of policy, however, patients preferences for different processes associated with care may be borne out through alternative means such as treatment adherence or compliance. Future research should consider this further, to ensure no double-counting of the impact of non health effects.

Conclusions

This review aimed to identify published studies which provide empirical measures of process utility and use methodology which allow them to be

incorporated into QALY calculations, and so included within cost-utility analyses. The studies identified demonstrate that the process of health care has been considered in relation to what enters an individual's utility function, and the evidence to date does support its existence not only in the administration of drugs, but also in approaches to screening and preventative care. This review indicates that patients place value on certain aspects of the process of receiving health care.

This review has also highlighted areas for further research including a comparative study of alternative methods for the measurement of process utility, the need for testing the validity of results through psychometric approaches and comparative studies with other patient reported measures, the "bolting-on" of a process domain to an existing generic PRM using already validated measures of patient experience and consideration of patients' adaptation to a health-state in the measurement of experienced utility.

When considering the number of economic evaluations which are published in relation to the number identified from this search it can be assumed that this aspect of an individual's utility function is rarely included in cost-utility analyses, and therefore many of these studies may be presenting results which do not truly reflect the impact on patients' health related quality of life and therefore are not an accurate estimation of cost-effectiveness.

Health economists therefore, need to acknowledge that patients attach value to process aspects and other non-health outcomes of health care interventions. Whilst is it acknowledged that there is no sound

methodological framework available at present for incorporating “process” into the QALY. One approach may be to use alternative methodology e.g. discrete choice experiments, which would allow a measure of “process” (although non-compatible with QALY’s), and to use the findings in combination with QALY-outcomes. However, recognising the importance of process and non-health outcomes also recognises the need to develop these QALY friendly approaches further.

The implication of the findings of this review is that health economists and decision makers need to acknowledge that aspects of the process of receiving health care may enter into a patient’s utility function. Whilst acknowledging the limitations of the current evidence base, consideration should be given on a study by study basis as to whether this should be incorporated into a cost-utility analysis. If results are produced, with and without the incorporation of process utility, decision makers can assess the sensitivity of a decision to its inclusion. This would allow comparability to be maintained via the base case analysis, whilst allowing methods to be developed that are of value to decision makers.

References

1. Oliver A. Prioritizing Health Care: Is “Health” Always an Appropriate Maximand? Med Decis Making 2004; 24: 272.

2. Sussex J, Towse A, Devlin N Operationalizing value-based pricing of medicines. A taxonomy of approaches. PharmacoEconomics (2013) 31:1-10.

3. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull. 2010; 96: 5–21.

4. Werner B.F. Brouwer, Anthony J. Culyer, N. Job A. van Exel, et al. Welfarism vs. extra-welfarism 2008, Journal of Health Economics 2008; 27 (2): 325–338

5. Gerard and Mooney. QALY league tables: handle with care. Health Economics 1993; 2: 59–64

6. Mooney G. Beyond health outcomes: The benefits of health care. Health Care Analysis. 1998; 6 (2): 99-1

7. McAlister D. Putting health economics into quality. Public Money & Management, 1994; 14 (2): 15-22

8. Donaldson C, Shackley P. Does "process utility" exist? A case study of willingness to pay for laproscopic cholcystectomy. Soc Sci Med 1997; 44: 699-707

9. Swan J, Sainfort F. Process Utility for Imaging in cerebrovascular disease. Acad Radiol 2003; Mar: 10(3): 266-74

10. Opmeer BC, de Borgie CA, Mol BW, Bossuyt PM. Assessing Preferences Regarding Healthcare Interventions that Involve Non-Health Outcomes: An Overview of Clinical Studies. Patient 2010;3(1):1-10.

11. Nord E, Pinto-Prades JL, Richardson J, et al. Incorporating societal concerns for fairness in numerical valuations of health programmes. Health Econ 1999;8:25–39

12. Tsuchiya A. QALYs and ageism: philosophical theories and age weighting. Health Econ 2000;9:57–68

13. Dolan, Paul and Tsuchiya , Aki (2010) Determining the parameters in social welfare function using stated preference data: an application to health. Applied economics, pp. 1466-4283

14. Drummond MF, Sculpher MJ,Torrance GW, O'Brien BJ, Stoddart Methods for the Economic Evaluation of Health Care Programmes. Third Edition. Oxford, UK.

15. Lloyd A, Nafees B, Bernett AH, et al. Willingness to Pay for Improvements in Chronic Long-Acting Insulin Therapy in Individuals With Type 1 or Type 2 Diabetes Mellitus. Clinical Therapeutics 2011; 33 (9): 1258-1267

16. Cook J, Richardson J, Street A. A Cost utility analysis of treatment options for gall-stone disease: methodological issues and results. Health Economics 1994; 3: 157-168

17. Schmier JK, Palmer CS, Flood EM, et al. Utility assessments of opioid treatment for chronic pain. Pain Med 2002; 3 (3): 218-30.

18. Osborne RH, De Abreu Lourenço R, Dalton A, Houltram J, et al. Quality of Life Related to Oral versus Subcutaneous Iron Chelation: A Time Trade-off Study; Value Health 2007; 10 (6) 51-6.

19. Osbourne RH, Dalton A, Hertel J, Schrover R, Smith DK. Health-related quality of life advantage of long-acting injectable antipsychotic treatment for schizophrenia: a time-trade-off study. Health and Quality of Life Outcomes 2012, 10:35.

20. Kauf TL, Roskell N, Shearer A, et al. A predictive model of health state utilities for HIV patients in the modern era of highly active antiretroviral therapy. Value Health 2008; 11 (7):1144-53.

21. Chancellor J, Aballéa S, Lawrence A, et al. Preferences of patients with diabetes mellitus for inhaled versus injectable insulin regimens. Pharmacoeconomics 2008; 26 (3):217-34.

22. Polster M, Zanutto E, McDonald S, et al. A comparison of preferences for two GLP-1 products--liraglutide and exenatide--for the treatment of type 2 diabetes. J Med Econ 2010; 13 (4):655-61.

23. Boye KS, Matza LS, Walter KN, et al. Utilities and disutilities for attributes of injectable treatments for type 2 diabetes. Eur J Health Econ. 2011 Jun;12 (3):219-30.

24. Birch S, Melnikow J, Kuppermann M. Conservative versus aggressive follow up of mildly abnormal Pap smears: testing for process utility. Health Econ. 2003; 12 (10):879-84

25. Howard K, Salkeld G, McCaffery K, et al. HPV triage testing or repeat Pap smear for the management of atypical squamous cells (ASCUS) on Pap smear: is there evidence of process utility? Health Econ. 2008; 17 (5):593-605

26. Cairns J, Shackley P, Hundley V. Decision Making with Respect to Diagnostic Testing: A method of Valuing the Benefits of Antenatal Screening. Med Decis Making 1996; 16: 61.

27. Swan JS, Fryback DG, Lawrence WF, et al. A Time-trade-off method for cost effectiveness models applied to radiology. Med Decis Making 2000; 20:79.

28. Swan JS, Lawrence WF, Roy J. Process utility in breast biopsy. Med Decis Making 2006; 26 (4):347-59

29. Salkeld G, Quine S, Cameron ID. What constitutes success in preventive health care? A case study in assessing the benefits of hip protectors. Soc Sci Med 2004; 59 (8):1593-601

30. Torrance GW. Measurement of health state utilities for economic appraisal. Journal of Health Economics 1986; 5: 1-30.

31. Walters SJ, Brazier JE. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. Health Qual Life Outcomes 2003;1:4.

32. Stouthard MEA, Essink-Bot ML, Bonsel GJ, Barendregt JJ, Kramer PG, vande Water HPA, Gunning-Schepers LJ, van der Maas PJ (1997). Disability Weights for Diseases in the Netherlands. Erasmus University, Rotterdam.

33. Boyd NF, Sutherland HJ, Heasman DL et al. Whose utilities for decision analysis? Med Dec Making 1990;10:58

34. Gold MR, Siegel JE, Russell LB, et al., editors. Cost-effectiveness in health and Medicine. Oxford: Oxford University Press 1996.

35. Brazier J, Akehurst R, Brennan A, Dolan P, Claxton K, McCabe C, Sculpher M, Tsuchyia A. Should patients have a greater role in valuing health states? Appl Health Econ Health Policy. 2005;4(4):201-8.)

36. Gafni A, Zylak CJ. Ionic versus non-ionic contrast media: a burden or a bargain? Can Med Assoc J 1990; 143 (6):475-478.

37. Brazier, J.E., Rowen, D. NICE DSU Technical Support Document 11: Alternatives to EQ-5D for generating health state utility values. 20011. Available from http://www.nicedsu.org.uk [Accessed 18 Sept 2012]

38. Krabbe P, Stouthard M, Essink-Bot M, et al. The effect of adding a cognitive dimension to the EuroQol Multiattribute Health-Status Classification System. J clin Epidemiol 1999; 52 (4): 293-301

39. Yang Y, Brazier J, Tsuchiya. The effect of adding a 'sleep' dimension to EQ-5D. Health Economists' Group meeting January 2008.

40. Brazier J, Rowen D, Tsuchiya A, Yang Y, Young T. What a pain: adding a generic dimension to a condition-specific preference-based measure. HESG Abstract, 2010.

41. Steine S, Finset A, Laerum E A new, brief questionnaire (PEQ) developed in primary health care for measuring patients' experience of interaction, emotion and consultation outcome. Family Practice 2001; 18:410-418.

42. Baron. Baises in the quantitative measurement of values for public decisions. Psychological Bulletin 1997

43. Dolan and Kahneman. Interpretations Of Utility And Their Implications For The Valuation Of Health. Economic Journal Volume 118, Issue 525, pages 215–234, January 2008

Appendix 11: Response to Letter to Editor: Capturing disutility from waiting time.

This response was written in reply to a Letter to the Editor: Gandjour, G. (2014) Letter to the Editor: Capturing Disutility from Waiting Time. *Pharmacoeconomics*. 32(4):423-424.

The final publication is: Brennan, V. and Dixon, S. (2014) Response to letter to editor: Capturing disutility from waiting time. Pharmacoeconomics. 2014 Apr;32(4):421-2, and is available at Springer via http://dx.doi.org/ 10.1007/s40273-013-0127-5.

Re: Manuscript Number PECA-D-13-00299.

Many thanks for the letter which suggests the use of the HYE to capture disutility associated with waiting for diagnostic tests, particularly for patients with a high probability of a poor diagnostic outcome. The author rightly points out that this approach was not used within any of the studies identified by the review, and argues that the methodologies which have been used do not adequately capture the disutility associated with any anxiety related to the uncertainty of the future test/procedure.

We would initially like to point out that the focus of the review and search strategy were developed specifically to identify those studies which provided an empirical measure of process utility which could be incorporated into the QALY calculation. This topic was identified as particularly important due to the wide-spread use of the QALY, driven predominantly by the requirements dictated by regulatory bodies such as NICE. As such, their point (which relates to the most appropriate method to capture utility irrespective of the analytical framework), is a different question to that which we addressed (which relates to those methods that are compatible with the conventional QALY framework).

However, this point aside, the letter argues that the requirement for a constant health state in QALY estimation (as opposed to a health profile valuation for HYE calculations), renders it unable to capture fully disutility related to anxiousness caused by uncertainty whilst waiting for diagnostic tests, and suggests this is particularly relevant where patients have a high probability of a poor diagnostic outcome. This is not entirely correct, as the study by Cook et al (1994) which was identified in our review estimated disutility associated with an impending operation within an ex ante QALY perspective. Although it should be noted that Cook and colleagues needed to dispose of the ex poste perspective of the conventional QALY model in order to capture disutility in this way.

The author also suggests that the requirement for QALY estimation to be based on a constant health state (as opposed to a health profile) renders it unable to capture disutility associated with the uncertainty and anxiousness experienced by patients awaiting diagnostic tests, and therefore suggest the use of the HYE or a variant of this proposed by Gandjour (2008) as more appropriate alternatives. The author also recognises that although this may incorporate uncertainty about future health through the use of probability distributions over future outcomes in a health profile, the approach still has the disadvantages which have prevented its uptake in the past associated with the complexities of valuing complicated health profiles (Towers et al, 2005).

As such, whilst we agree that HYEs offer an alternative, we feel that their superiority remains unproven. We also maintain our view that the prominence of the QALY within health care decision making suggests that the focus of future research should be on methods which are compatible with the conventional QALY model. Our review is the first step in any such programme of research.

References.

1. Cook, J. Richardson, J. Street, A. A cost utility analysis of treatment options for gallstone disease: methodological issues and results. Health Econ. 1994;3:157–68.

2. Gandjour A. Incorporating feelings related to the uncertainty about future health in utility measurement. Health Economics 2008 Oct;17(10):1207-13.

3. Towers I, Spencer A, Brazier J. Healthy year equivalents versus quality-adjusted life years: the debate continues. Expert Review of Pharmacoeconomics and Outcomes Research 2005;5(3):245–254.

Appendix 12: Results of two independent sample t-tests.

Exploratory analysis was performed on the PEQ and CSQ-8 data from the ePAQ trial. Parametric, two independent sample t-tests were performed as the data were found to be normally distributed based on a comparison of mean and median domain scores (PEQ) and total scores (CSQ-8) shown in Table 36. The presence of a significant difference between arms as was assumed to reflect a difference induced by the process of care, as the key difference between the trial arms was the mode of consultation (face-to-face versus telephone).

Table 36. **Results of two independent samples t test for PEQ and CSQ items, domains and total scores, for each trial group**

	Control				Intervention					95% CI		
	N	Medi-an	Mean	SD	N	Medi-an	Mean	SD	Mean Diff	Lower	Upper	P-value
PEQ Dimension												
Short term outcome	70	3	2.93	1.09	68	3.00	2.87	1.16	-.05	-.43	.32	.78
Communications	85	4	4.00	.75	85	4.00	4.23	.63	.37	.16	.58	.001
Emotions	85	4.5	4.69	1.27	81	5.25	5.32	1.18	.63	.26	1.01	.001
Barriers	87	4.00	4.12	.68	85	4.25	4.35	.58	.23	.04	.42	.018
PEQ Question												
Consultation Doctor	89	2	2.10	1.01	87	1	1.28	.64	-.83	-1.08	-.57	.000
Know what to do to reduce pelvic floor problems	86	3	2.79	1.28	80	3	2.7	1.38	-.09	-.50	.32	.662
Know what to expect from now on	86	4	3.41	1.25	85	4	3.48	1.23	.08	-.30	.45	.692

	Control				Intervention					95% CI		
	N	Medi-an	Mean	SD	N	Medi-an	Mean	SD	Mean Diff	Lower	Upper	P-value
Able to handle pelvic floor problems differ-ently	84	3	2.74	1.25	77	2	2.36	1.32	-.37	-.78	.03	.067
Lead to fewer or pre-vent pelvic floor prob-lems	71	3	2.73	1.24	70	3	2.71	1.42	-.02	-.46	.43	.936
Had a good talk	87	4	3.85	.84	85	4	4.28	.68	.43	.20	.66	.000
Felt reassured	87	4	3.71	.82	86	4	4.08	.79	.37	.13	.61	.003
Clinician understood what was on my mind	88	4	3.80	.89	86	4	4.27	.69	.47	.24	.71	.000
Felt was taken care of	87	4	4.02	.73	87	4	4.26	.69	.24	.03	.46	.026
Bit difficult to connect with clinician	88	4	3.98	.96	85	4	4.25	.79	.27	.01	.53	.045

	Control				Intervention					95% CI		
	N	Medi-an	Mean	SD	N	Medi-an	Mean	SD	Mean Diff	Lower	Upper	P-value
Too much time spent on small talk	87	4	4.26	.66	87	4	4.37	.65	.16	-.019	.34	.079
Bit difficult to ask questions	88	4	3.91	.98	87	4	4.28	.76	.38	.12	.64	.005
Important decisions made over head	86	4	4.30	.75	87	4	4.37	.65	.07	-.14	.28	.494
Relieved - Worried	86	5	4.80	1.49	86	6	5.50	1.24	.70	.29	1.11	.001
Cheerful - Sad	85	4	4.55	1.52	81	6	5.37	1.47	.82	.36	1.3	.001
Strengthened - Worn out	85	4	4.71	1.20	83	5	5.12	1.40	.42	.02	.81	.041
Relaxed - Tense	85	5	4.71	1.60	84	5	5.25	1.42	.54	.09	1.00	.020

	Control				Intervention					95% CI		
	N	**Medi-an**	**Mean**	**SD**	**N**	**Medi-an**	**Mean**	**SD**	**Mean Diff**	**Lower**	**Upper**	**P-value**
CSQ												
Total score	79	28.00	27.16	3.79	75	28	26.85	4.48	-.64	-1.63	1.01	.285
CSQ Items												
CSQ: Rate service received	89	3	3.39	.63	87	4	3.44	.64	.04	-.15	.23	.651
CSQ: Kind of service wanted	88	3	3.34	.59	85	3	3.35	.65	.01	-.17	.20	.371
CSQ: Extent service met needs	88	3	3.17	.65	83	3	3.11	.78	-.06	-.28	.15	.572
CSQ: Would recom-mend service	87	4	3.52	.53	85	4	3.46	.63	-.06	-.23	.17	.509
CSQ: Satisfied with amount of help	88	3.5	3.43	.64	86	3	3.38	.71	-.05	-.25	.15	.638

	Control				Intervention					95% CI		
	N	**Medi-an**	**Mean**	**SD**	**N**	**Medi-an**	**Mean**	**SD**	**Mean Diff**	**Lower**	**Upper**	**P-value**
CSQ: Services helped to deal more effec-tively with problems	84	3	3.12	.70	78	3	3.06	.69	-.06	-.27	.16	.616
CSQ: Overall, satis-fied with service	88	3	3.36	.63	87	3	3.32	.71	-.04	-.24	.16	.680
CSQ: If were to seek help again, would come back to services	87	4	3.55	.55	85	4	3.47	.61	-.08	-.26	.09	.642

Appendix 13: Principal Component Analysis

PCA was performed using SPSS on the PEQ and CSQ-8 questionnaires. As illustrated in Table 37, the Kaiser-Meyer-Olkin (KMO) measure verified the sampling adequacy for the analysis, KMO = .910. All of the KMO values for individual items, which can be seen highlighted on the diagonal of the anti-image correlation matrix were >.830 which is well above the acceptable limit of .5 (Field, 2009). Bartlett's test of sphericity X2 (276) = 2241.357, p<.001, indicated that correlations between items were sufficiently large for PCA.

Table 37. KMO test and Bartlett's test of Sphericity.

KMO Test	
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.910
Bartlett's Test of Sphericity	
Approx. Chi-Square	2241.357
df	276.000
Sig.	.000

Table 38 shows the KMO values for individual variables, these are the diagonal elements of the anti-image correlation matrix (in bold). These values are all above

the minimum .5 value. Therefore all variables were included in the analysis. Bartlett's test of sphericity X2 (276) = 2241.357, p<.001, was also significant, this indicates there are some relationships between the variables we hope to include in the analysis.

Table 38. Anti-image matrices

	CSQ: Rate service received	CSQ: Kind of service wanted	CSQ: Extent service met needs	CSQ: Would recommend service	CSQ: Satisfied with amount of help	CSQ: Services helped to deal more effectively with problems	CSQ: Overall, satisfied with service	CSQ: If were to seek help again, would come back to services
CSQ: Rate service received	**.924**[a]	-.300	-.121	-.091	.066	.040	.056	-.155
CSQ: Kind of service wanted	-.300	**.929**[a]	-.253	-.080	-.064	.045	-.070	-.087
CSQ: Extent service met needs	-.121	-.253	**.942**[a]	-.044	.020	-.296	-.217	-.006
CSQ: Would recommend service	-.091	-.080	-.044	**.876**[a]	-.047	-.025	-.147	-.452
CSQ: Satisfied with amount of help	.066	-.064	.020	-.047	**.912**[a]	-.043	-.242	-.525
CSQ: Services helped to deal more effectively with problems	.040	.045	-.296	-.025	-.043	**.943**[a]	-.188	-.081
CSQ: Overall, satisfied with service	.056	-.070	-.217	-.147	-.242	-.188	**.946**[a]	-.044
CSQ: If were to seek help again, would come	-.155	-.087	-.006	-.452	-.525	-.081	-.044	**.871**[a]

	CSQ: Rate service received	CSQ: Kind of service wanted	CSQ: Extent service met needs	CSQ: Would recommend service	CSQ: Satisfied with amount of help	CSQ: Services helped to deal more effectively with problems	CSQ: Overall, satisfied with service	CSQ: If were to seek help again, would come back to services
back to services								
PEQ: Know what to do to reduce pelvic floor problems	-.095	.037	.070	-.034	.030	-.064	-.006	-.029
PEQ: Know what to expect from now on	.091	-.118	-.022	.187	-.105	-.042	-.044	.033
PEQ: Able to handle pelvic floor problems differently	-.036	.043	.008	-.212	-.087	-.152	.061	.207
PEQ: Lead to fewer or prevent pelvic floor problems	.132	-.006	-.084	.037	.094	-.096	.118	-.122
PEQ: Had a good talk	-.081	-.084	.097	-.250	-.185	.104	.025	.325
PEQ: Felt reassured	-.053	.014	-.036	.144	-.008	-.046	-.034	-.051
PEQ: Clinician understood what was on my mind	-.174	.139	.063	.295	.104	-.074	-.181	-.199

	CSQ: Rate service received	CSQ: Kind of service wanted	CSQ: Extent service met needs	CSQ: Would recommend service	CSQ: Satisfied with amount of help	CSQ: Services helped to deal more effectively with problems	CSQ: Overall, satisfied with service	CSQ: If were to seek help again, would come back to services
PEQ: Felt was taken care of	.007	-.091	-.008	.081	-.013	.065	.078	-.154
PEQ: Bit difficult to connect with clinician	.281	-.096	-.098	-.155	.005	.094	.094	-.033
PEQ: Too much time spent on small talk	-.137	-.048	.065	.155	.061	-.105	-.039	-.129
PEQ: Bit difficult to ask questions	.062	.145	-.158	-.130	.077	-.012	-.002	.027
PEQ: Important decisions made over head	-.189	.029	.007	.002	-.201	.150	-.219	.233
PEQ: Relieved - Worried	-.028	.015	-.072	.222	-.134	-.003	-.138	-.024
PEQ: Cheerful - Sad	-.119	.318	-.173	-.180	-.279	.041	.070	.271
PEQ: Strengthened - Worn out	.137	-.185	.113	-.220	.231	-.154	.170	-.076
PEQ: Relaxed - Tense	-.009	-.171	.077	.069	.095	.041	-.064	-.084

	PEQ: Know what to do to reduce pelvic floor problems	PEQ: Know what to expect from now on	PEQ: Able to handle pelvic floor problems differently	PEQ: Lead to fewer or prevent pelvic floor problems	PEQ: Had a good talk	PEQ: Felt reassured	PEQ: Clinician understood what was on my mind	PEQ: Felt was taken care of
CSQ: Rate service received	-.095	.091	-.036	.132	-.081	-.053	-.174	.007
CSQ: Kind of service wanted	.037	-.118	.043	-.006	-.084	.014	.139	-.091
CSQ: Extent service met needs	.070	-.022	.008	-.084	.097	-.036	.063	-.008
CSQ: Would recommend service	-.034	.187	-.212	.037	-.250	.144	.295	.081
CSQ: Satisfied with amount of help	.030	-.105	-.087	.094	-.185	-.008	.104	-.013
CSQ: Services helped to deal more effectively with problems	-.064	-.042	-.152	-.096	.104	-.046	-.074	.065
CSQ: Overall, satisfied with service	-.006	-.044	.061	.118	.025	-.034	-.181	.078
CSQ: If were to seek help again,	-.029	.033	.207	-.122	.325	-.051	-.199	-.154

	PEQ: Know what to do to reduce pelvic floor problems	PEQ: Know what to expect from now on	PEQ: Able to handle pelvic floor problems differently	PEQ: Lead to fewer or prevent pelvic floor problems	PEQ: Had a good talk	PEQ: Felt reassured	PEQ: Clinician understood what was on my mind	PEQ: Felt was taken care of
would come back to services								
PEQ:Know what to do to reduce pelvic floor problems	**.872[a]**	-.190	-.382	-.340	-.051	-.121	-.054	.115
PEQ: Know what to expect from now on	-.190	**.911[a]**	-.094	-.144	.034	-.172	-.102	-.039
PEQ: Able to handle pelvic floor problems differently	-.382	-.094	**.830[a]**	-.334	.095	.054	.019	-.184
PEQ: Lead to fewer or prevent pelvic floor problems	-.340	-.144	-.334	**.842[a]**	-.016	.041	-.009	.006
PEQ: Had a good talk	-.051	.034	.095	-.016	**.876[a]**	-.418	-.157	-.369
PEQ: Felt reassured	-.121	-.172	.054	.041	-.418	**.925[a]**	-.060	-.113
PEQ: Clinician understood what was on my mind	-.054	-.102	.019	-.009	-.157	-.060	**.934[a]**	-.335
PEQ: Felt was taken care of	.115	-.039	-.184	.006	-.369	-.113	-.335	**.936[a]**

	PEQ: Know what to do to reduce pelvic floor problems	PEQ: Know what to expect from now on	PEQ: Able to handle pelvic floor problems differently	PEQ: Lead to fewer or prevent pelvic floor problems	PEQ: Had a good talk	PEQ: Felt reassured	PEQ: Clinician understood what was on my mind	PEQ: Felt was taken care of
PEQ: Bit difficult to connect with clinician	.046	-.033	-.055	.027	-.094	-.062	-.157	-.069
PEQ: Too much time spent on small talk	.172	.257	-.033	-.158	-.230	-.033	-.008	.133
PEQ: Bit difficult to ask questions	.019	-.081	.121	.035	-.048	-.071	-.115	-.086
PEQ: Important decisions made over head	-.001	.067	-.019	-.169	.184	.029	-.103	-.014
PEQ: Relieved - Worried	.016	.210	-.054	-.135	-.066	-.135	.025	.056
PEQ: Cheerful - Sad	.025	-.207	.160	-.051	.151	-.118	.115	-.171
PEQ: Strengthened - Worn out	.043	-.051	.015	.130	.192	-.312	-.182	.066
PEQ: Relaxed - Tense	-.116	-.091	-.097	.111	-.191	.376	-.017	.071

	PEQ: Bit difficult to connect with clinician	PEQ: Too much time spent on small talk	PEQ: Bit difficult to ask questions	PEQ: Important decisions made over head	PEQ: Relieved - Worried	PEQ: Cheerful - Sad	PEQ: Strengthened - Worn out	PEQ: Relaxed - Tense
CSQ: Rate service received	.281	-.137	.062	-.189	-.028	-.119	.137	-.009
CSQ: Kind of service wanted	-.096	-.048	.145	.029	.015	.318	-.185	-.171
CSQ: Extent service met needs	-.098	.065	-.158	.007	-.072	-.173	.113	.077
CSQ: Would recommend service	-.155	.155	-.130	.002	.222	-.180	-.220	.069
CSQ: Satisfied with amount of help	.005	.061	.077	-.201	-.134	-.279	.231	.095
CSQ: Services helped to deal more effectively with problems	.094	-.105	-.012	.150	-.003	.041	-.154	.041
CSQ: Overall, satisfied with service	.094	-.039	-.002	-.219	-.138	.070	.170	-.064
CSQ: If were to seek help again, would come back to services	-.033	-.129	.027	.233	-.024	.271	-.076	-.084
PEQ: Know what to do to reduce	.046	.172	.019	-.001	.016	.025	.043	-.116

	PEQ: Know what to do to reduce pelvic floor problems	PEQ: Know what to expect from now on	PEQ: Able to handle pelvic floor problems differently	PEQ: Lead to fewer or prevent pelvic floor problems	PEQ: Had a good talk	PEQ: Felt reassured	PEQ: Clinician understood what was on my mind	PEQ: Felt was taken care of
pelvic floor problems								
PEQ: Know what to expect from now on	-.033	.257	-.081	.067	.210	-.207	-.051	-.091
PEQ: Able to handle pelvic floor problems differently	-.055	-.033	.121	-.019	-.054	.160	.015	-.097
PEQ: Lead to fewer or prevent pelvic floor problems	.027	-.158	.035	-.169	-.135	-.051	.130	.111
PEQ: Had a good talk	-.094	-.230	-.048	.184	-.066	.151	.192	-.191
PEQ: Felt reassured	-.062	-.033	-.071	.029	-.135	-.118	-.312	.376
PEQ: Clinician understood what was on my mind	-.157	-.008	-.115	-.103	.025	.115	-.182	-.017
PEQ: Felt was taken care of	-.069	.133	-.086	-.014	.056	-.171	.066	.071

	PEQ: Know what to do to reduce pelvic floor problems	PEQ: Know what to expect from now on	PEQ: Able to handle pelvic floor problems differently	PEQ: Lead to fewer or prevent pelvic floor problems	PEQ: Had a good talk	PEQ: Felt reassured	PEQ: Clinician understood what was on my mind	PEQ: Felt was taken care of
PEQ: Bit difficult to connect with clinician	**.940[a]**	-.255	-.060	-.304	-.006	.070	.039	-.114
PEQ: Too much time spent on small talk	-.255	**.903[a]**	-.125	-.082	.127	-.176	-.005	-.028
PEQ: Bit difficult to ask questions	-.060	-.125	**.958[a]**	-.218	-.050	.006	.066	-.159
PEQ: Important decisions made over head	-.304	-.082	-.218	**.881[a]**	.170	-.016	-.372	.160
PEQ: Relieved - Worried	-.006	.127	-.050	.170	**.925[a]**	-.157	-.326	-.343
PEQ: Cheerful - Sad	.070	-.176	.006	-.016	-.157	**.883[a]**	-.234	-.259
PEQ: Strengthened - Worn out	.039	-.005	.066	-.372	-.326	-.234	**.886[a]**	-.312
PEQ: Relaxed - Tense	-.114	-.028	-.159	.160	-.343	-.259	-.312	**.893[a]**

As described by Field, (2009), only those factors which are deemed statistically important are retained in a PCA. To determine the importance of the vector the magnitude of the associated eigenvalue is reviewed and based on Kaiser's criterion those factors with eigenvectors greater than 1 are retained. The results are shown in the main body of the thesis. Four components had eigenvalues over Kaiser's criterion of 1 and in combination explained 71.55% of the variance. This criterion is accurate where there are less than 30 variables and communalities after extraction are greater than 0.7 or when the sample size exceeds 250 and the average communality is greater than 0.6. In this study, there were 24 variables, and the mean communality was .715 (Table 39).

Table 39. **Communalities.**

	Initial	Extraction
CSQ: Rate service received	1.000	.578
CSQ: Kind of service wanted	1.000	.676
CSQ: Extent service met needs	1.000	.664
CSQ: Would recommend service	1.000	.733
CSQ: Satisfied with amount of help	1.000	.770
CSQ: Services helped to deal more effectively with problems	1.000	.648
CSQ: Overall, satisfied with service	1.000	.732
CSQ: If were to seek help again, would come back to services	1.000	.832
PEQ:Know what to do to reduce pelvic floor problems	1.000	.795

	Initial	Extraction
PEQ:Know what to expect from now on	1.000	.626
PEQ:Able to handle pelvic floor problems differently	1.000	.762
PEQ:Lead to fewer or prevent pelvic floor problems	1.000	.738
PEQ:Had a good talk	1.000	.771
PEQ:Felt reassured	1.000	.778
PEQ:Clinician understood what was on my mind	1.000	.792
PEQ:Felt was taken care of	1.000	.789
PEQ:Bit difficult to connect with clinician	1.000	.681
PEQ:Too much time spent on small talk	1.000	.541
PEQ:Bit difficult to ask questions	1.000	.624
PEQ:Important decisions made over head	1.000	.498
PEQ:Relieved - Worried	1.000	.792
PEQ:Cheerful - Sad	1.000	.751
PEQ:Strengthened - Worn out	1.000	.834
PEQ:Relaxed - Tense	1.000	.769

Appendix 14: Process Utility Pilot questionnaire booklet.

Respondent ID _____

Introduction

Thank you for agreeing to take part in this survey. This is a survey for the School of Health and Related Research at the University of Sheffield.

Different people experience different states of health. Some people's health is better than others. We are interested in comparing different states of health and measuring how good or bad they are.

This questionnaire is part of a pilot study to determine the preferred descriptions of health states.

We will ask you to think about a set of 3 different states of health and tell us how good or bad you think they are.

We will then ask you to think about another set of 3 different health states which are described differently.

We ask you which set of descriptions you prefer.

We will then ask you to think about 5 scenarios describing your feelings immediately after a consultation with the doctor, and tell us if you would consider the communication experience to be very good, good, average, poor or very poor.

All of your responses will be treated as confidential, and all analysis will be carried out anonymously.

We are interested in people's views, and there are no right or wrong answers. Please tell us what you think.

Please read this page carefully before going further.

The following pages contain descriptions of different states of health. Each one is in a box.

For each box, imagine that you have to live in that state for one week.

How good or bad do you think each state of health is compared to the others?

Please give each state of health a score of between 0 and 100, where 0 = the worst state of health you can imagine and 100= the best state of health you can imagine.

Write the score in the box at the bottom of each health state.

The health states on the following pages are described using

Method A.

On a scale between 0 and 100 like the one shown, how would you SCORE this health state?

Your health does not limit you in moderate activities

You have no problems with your work of other regular daily activities as a result of your physical health or any emotional problems

Your health limits you social activities none of the time

You have no pain

You feel tense or downhearted and low none of the time

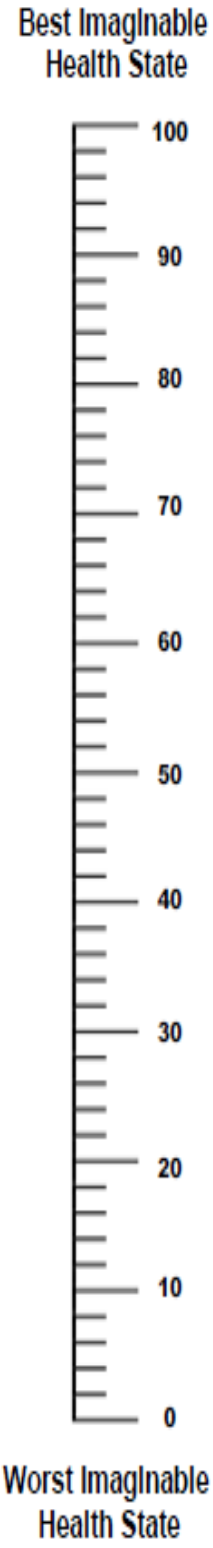
You have a lot of energy all of the time

You have just had a consultation with the doctor. When asked to evaluate the communication experience your response was:

The clinician understood what was on my mind

Agree completely Agree So-so Disagree Disagree completely

SCORE = ☐


Best Imaginable
Health State
100
90
80
70
60
50
40
30
20
10
0
Worst Imaginable
Health State

On a scale between 0 and 100 like the one shown, how would you SCORE this health state?

Your health limits you a little in moderate activities

You accomplish less than you would like as a result of emotional problems.

Your health limits your social activities some of the time

You have pain that interferes with your normal work (both outside the home and housework) moderately.

You feel tense or downhearted and low some of the time

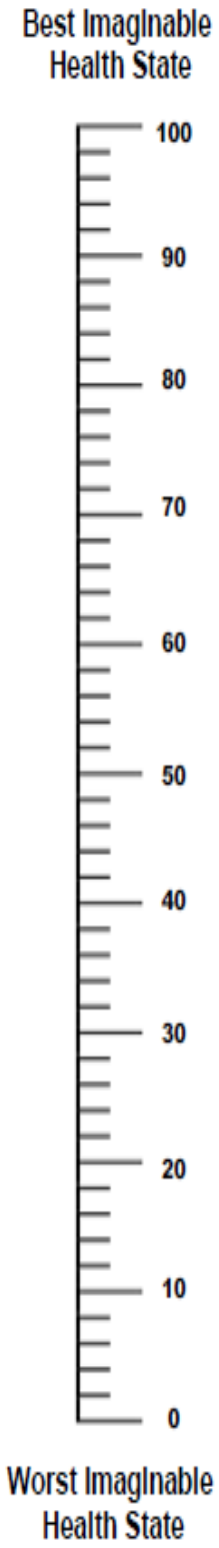
You have a lot of energy some of the time

You have just had a consultation with the doctor. When asked to evaluate the communication experience your response was:

The clinician understood what was on my mind

Agree completely Agree (So-so) Disagree Disagree completely

SCORE = [ ]


Best Imaginable
Health State
100
90
80
70
60
50
40
30
20
10
0
Worst Imaginable
Health State

On a scale between 0 and 100 like the one shown, how would you SCORE this health state?

Your health limits you a lot in moderate activities

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

You health limits you social activities all of the time

You have pain that interferes with your normal work (both outside the home and housework) extremely

You feel tense or downhearted and low all of the time

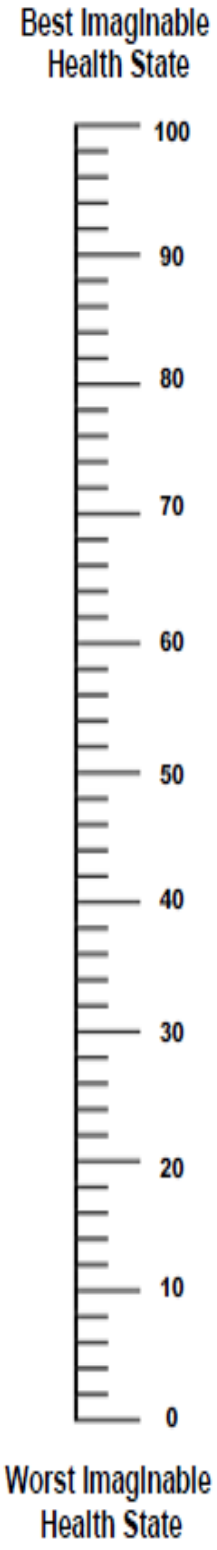
You have a lot of energy none of the time

You have just had a consultation with the doctor. When asked evaluate the communication experience your response was:

The clinician understood what was on my mind

Agree completely Agree So-so Disagree (Disagree completely)

SCORE = [ ]


Best Imaginable
Health State
100
90
80
70
60
50
40
30
20
10
0
Worst Imaginable
Health State

This page has been intentionally left blank

The health states on the following pages are described using

Method B.

On a scale between 0 and 100 like the one shown, how would you SCORE this health state?

Your health does not limit you in moderate activities

You have no problems with your work of other regular daily activities as a result of your physical health or any emotional problems

Your health limits you social activities none of the time

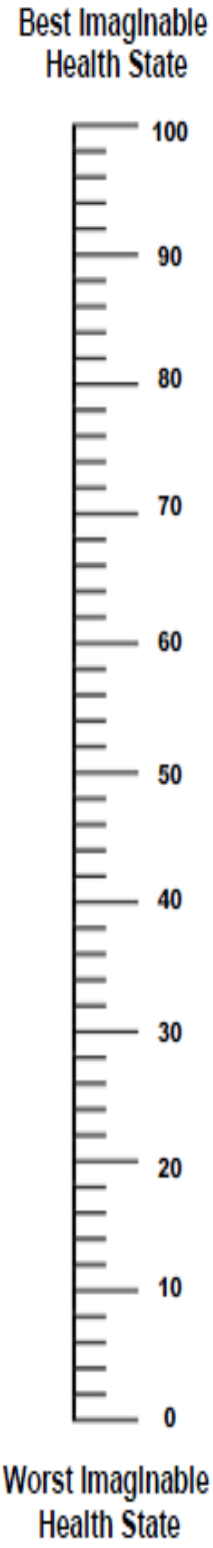
You have no pain

You feel tense or downhearted and low none of the time

You have a lot of energy all of the time

You have just had a consultation with the doctor. You evaluated your communication experience as very good.

SCORE = ☐


Best Imaginable
Health State
100
90
80
70
60
50
40
30
20
10
0
Worst Imaginable
Health State

On a scale between 0 and 100 like the one shown, how would you SCORE this health state?

Your health limits you a little in moderate activities

You accomplish less than you would like as a result of emotional problems.

Your health limits your social activities some of the time

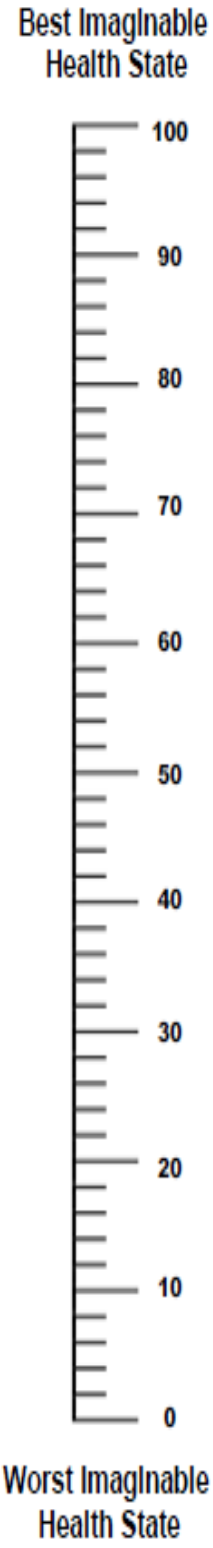
You have pain that interferes with your normal work (both outside the home and housework) moderately.

You feel tense or downhearted and low some of the time

You have a lot of energy some of the time

You have just had a consultation with the doctor. You evaluated your communication experience as average.

SCORE = ☐


Best Imaginable
Health State
100
90
80
70
60
50
40
30
20
10
0
Worst Imaginable
Health State

On a scale between 0 and 100 like the one shown, how would you SCORE this health state?

Your health limits you a lot in moderate activities

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

You health limits you social activities all of the time

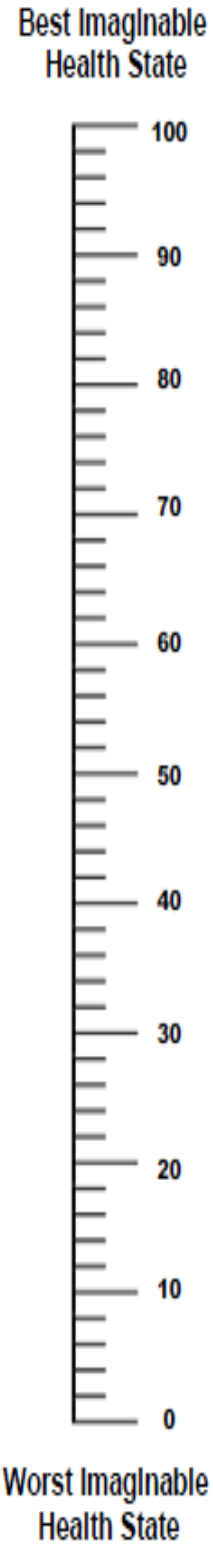
You have pain that interferes with your normal work (both outside the home and housework) extremely

You feel tense or downhearted and low all of the time

You have a lot of energy none of the time

You have just had a consultation with the doctor. You evaluated your communication experience as very poor.

SCORE = ☐


Best Imaginable
Health State
100
90
80
70
60
50
40
30
20
10
0
Worst Imaginable
Health State

Which method do you prefer?

By placing a tick in one box please indicate which method for describing health states you prefer.

I prefer Method A ☐

I prefer Method B ☐

Please give reasons for your choice.

Please read this page carefully before going further.

The following pages contain scenarios describing your feelings immediately after a consultation with the doctor.

For each scenario, please tell us if you would consider the consultation to be:

Very poor ☐

Poor ☐

Average ☐

Good ☐

Very good ☐

You have just had a consultation with your doctor. Immediately after the consultation your responses were:

We had a good talk

Agree completely Agree So-so Disagree Disagree completely

I felt reassured

Agree completely Agree So-so Disagree Disagree completely

The clinician understood what was on my mind

Agree completely Agree So-so Disagree Disagree completely

I felt I was taken care of

Agree completely Agree So-so Disagree Disagree completely

Would you consider this consultation to be?

Very poor ☐

Poor ☐

Average ☐

Good ☐

Very good ☐

You have just had a consultation with your doctor. Immediately after the consultation your responses were:

We had a good talk

Agree completely Agree So-so (Disagree) Disagree completely

I felt reassured

Agree completely Agree So-so (Disagree) Disagree completely

The clinician understood what was on my mind

Agree completely Agree So-so (Disagree) Disagree completely

I felt I was taken care of

Agree completely Agree So-so (Disagree) Disagree completely

Would you consider this consultation to be?

Very poor ☐

Poor ☐

Average ☐

Good ☐

Very good ☐

You have just had a consultation with your doctor. Immediately after the consultation your responses were:

We had a good talk

Agree completely Agree So-so Disagree (Disagree completely)

I felt reassured

Agree completely Agree So-so Disagree (Disagree completely)

The clinician understood what was on my mind

Agree completely Agree So-so Disagree (Disagree completely)

I felt I was taken care of

Agree completely Agree So-so (Disagree) Disagree completely

Would you consider this consultation to be?

Very poor ☐

Poor ☐

Average ☐

Good ☐

Very good ☐

You have just had a consultation with your doctor. Immediately after the consultation your responses were:

We had a good talk

Agree completely Agree (So-so) Disagree Disagree completely

I felt reassured

Agree completely Agree (So-so) Disagree Disagree completely

The clinician understood what was on my mind

Agree completely (Agree) So-so Disagree Disagree completely

I felt I was taken care of

Agree completely Agree (So-so) Disagree Disagree completely

Would you consider this consultation to be?

Very poor ☐

Poor ☐

Average ☐

Good ☐

Very good ☐

You have just had a consultation with your doctor. Immediately after the consultation your responses were:

We had a good talk

Agree completely Agree (So-so) Disagree Disagree completely

I felt reassured

Agree completely Agree (So-so) Disagree Disagree completely

The clinician understood what was on my mind

Agree completely (Agree) So-so Disagree Disagree completely

I felt I was taken care of

Agree completely Agree (So-so) Disagree Disagree completely

Would you consider this consultation to be?

Very poor ☐

Poor ☐

Average ☐

Good ☐

Very good ☐

Appendix 15: SSI Proposal

Proposal for University of Sheffield
Friday, February 24, 2012

General Population sample only

QUO-98468-42CC

Type of service: Sample Only
Project type: Standard
Dedupe: Not Known

Project Description

I am a PhD student at Sheffield University and as part of my studies I am undertaking a study which involves an online questionnaire. The questionnaire itself is being set up by a company called EpiGenesis and is not yet completely developed but will take approximately 20 minutes to complete. I am in the process of applying for grants to fund the study and therefore need quotes from Epigenysis and a panel company. Brendan Mulhern from Sheffield University gave me your contact details for a panel company he had worked with in the past.

For my study there will be 2 versions of the questionnaire and I am aiming to have 100 people complete each questionnaire. The sample needs to be representative of the general population. Each questionnaire will take no more than 20 minutes to complete

SAMPLE ONLY

200 completes

100% incidence

20 minute survey

Pricing

Total cost: £762.00

Summary Level	Completes	CPI [GBP]	Total Cost [GBP]
All Countries and Additional Services	200	£3.81	£762.00
		Total Price:	**£762.00**

Country	Completes	CPI [GBP]	Total Cost [GBP]
UK 100.0% 20min	200	£2.31	£462.00
		Sample Cost Total:	**£462.00**

Additional Services	Total Cost [GBP]
Project Management	£300.00
Additional Services Total:	**£300.00**

Purchase Now

I Have Questions

Pricing Terms

1. All prices are excluding tax.
2. Changes to the above specifications may affect the quoted price and timing. For example, if the actual incidence differs significantly from the estimated incidence, costs will be adjusted accordingly.
3. For sample-only projects, the client is responsible for installing quota stops. If the number of completes exceeds the target by more than 10%, all extra completes will be invoiced.

Thank you for the opportunity to describe the comprehensive global support SSI can provide for your sample needs.

We are confident we can convince you that SSI's vision and expanding global footprint position us as the best partner for the long term and for the continued success and growth of your research programs going forward. We believe that through this process we will demonstrate to you compelling reasons to choose SSI to support your needs in 2012.

Prepared for Victoria Brennan by Rosemary Greening, your dedicated and committed Business Development Manager.

I will accompany you from the feasibility and bidding phase all the way to the data delivery. My focus is on providing you with superior data and a superior experience in working with SSI.

How can you contact me

E-mail:
Rosemary_Greening@SurveySampling.com

Office phone:
+44 (0) 20 7421 1177

Mobile phone:
+44 07515 328 932

Rosemary Greening

Appendix 16: Epigenesys Proposal

epiGenesys
a University of Sheffield company

New Features: ScHARR Survey System

Feature	Description	Estimated days
Extension to rating questions to include a scenario.	Instead of just asking the participant to answer the rating question, this new feature will enable a scenario image to be illustrated above the question so that the rating question will be about the scenario itself.	2

Further costs

Liaison with panelling company: 1 days work x 400 per day = £400

The system has already been built to accommodate external participants recruited by a panelling company. However, the details of every survey are different so we need to spend time collaborating with the panelling company to make sure the setup of both systems is correct for your survey.

Totals

Development and liaison: 3 days work x 400 per day = £1200

Appendix 17: ScHARR Data Collection Fund Application

ScHARR

PGR Data Collection Fund Application

2012-2013

Name	Victoria Brennan
Registration number	070211960
Full time/Part time	Part time
Total amount requested (details overleaf)	£1,962
Brief description of the research activity i.e. survey, interviews, specialist software,	The planned study is an online health state valuation study which aims to explore the addition of a “process” domain to the SF-6D, using methodology which will allow the measure of process utility to be incorporated into the QALY calculation.

transcribing etc	The proposed study involves the development of an online health state valuation study, using the VAS, Two sets of health states will be developed.; the first comprised of SF-6D health states, and the second constructed similar to the first, except for the addition of a process dimension (SF-7D). Valuations of both sets of health states will be analysed to detect the effect of the additional dimension. Regression techniques will then be used to relate the findings from the valuation study back to a clinical trial which was the basis for the development of the process domain.
Are Ethics and Governance in place?	The online health state valuation study proposal will be submitted to ScHARR for ethical approval. However, ethics approval can not be sought until the online questionnaire has been developed, and the questionnaire can not be developed without the funding. Ethics are therefore not yet in place. Governance: The study proposal will also undergo ScHARR Governance procedures including scientific review. However, again this process can not be completed until funding has been secured and the online questionnaire developed. The funding is therefore required prior to ethics and governance being in place.
Reasons for the funding request e.g. how would this funding add value to your	I am requesting the funding for a health state valuation study which I am taking as part of my PhD studies. The study aims to explore the addition of a process utility domain to the SF-6D, and to derive a utility measure which can be incorporated into the QALY calculation. This valuation study was not included in my original

PhD	PhD plan, but was recommended at my upgrade and therefore is an unplanned expenditure. The finding is required to develop and host an online questionnaire which incorporates the online use of the VAS. Without the assistance of the funding it is highly unlikely that I will be able to recruit a sample size of greater than 50-60 as for the pilot study I was unable to reach a sample size of 30. A key paper by Krabbe et al., 1999 which used a similar approach reports using a convenience sample of 87 and discusses this as a key limitation to their study. This, and discussions with my supervisor Simon Dixon indicate that a sample size of nearer 200 is required, and therefore an online questionnaire is needed to facilitate this.
Describe the logistics involved i.e. sample size, location, transcription, travel, etc. These should match with the itemised costs and timescales on the page overleaf	The funding will allow me to develop the health state valuation questionnaire online. This will be performed through a company called epiGenesys who have been working with staff at ScHARR (Donna Rowan) to develop online questionnaires for health state preference studies. Once developed, the funding will allow me to work through a panel company to obtain a sample size of 200 people who are representative of the general population and therefore enhance the scientific rigour of the study.
Do you have any other sources of funding e.g. RTSG? If how much will you be	I have sought the funding through my current employer, who have said no as they are funding my attendance at a conference to present a poster related to my PhD work. I have also sought funding from Sheffield Teaching Hospital Trust, however, they only provide

using for this activity? If not, please explain why this cannot be used	funding for trust employees which I am not. I therefore have no other sources of funding.
Support from supervisor attached **(what the funding is for and how it fits in with your PhD / traning needs and and career aspirations that you may have).**	
Any additional info	In addition to the funding allowing significant value to my PhD, the study will be assisting Donna Rowen in testing the online application of health state valuation studies including the VAS, and would therefore be of additional benefit to both ScHARR and HEDS.

Please detail total expected expenditure below (in £ sterling)

Please itemise the costs in detail below	
epiGenisys: Liaison with paneling company	1 days work x £400/day = £400
epiGenysis: Extension to existing	2 days work x £400/day = £800

ScHARR survey system to include scenario rating question	
SSI panel company: to provide sample of 200, general population (Panel company suggested by Brendan Mulhern who has worked previously with ScHARR)	£762.00
TOTAL (in £ sterling)	1,962

Appendix 18: ScHARR Data Collection Fund Approval

Hi Victoria

I'm happy to approve your application. Please liaise with Brigid/Melanie to access funding.

Good luck,
Praveen

On 26 April 2012 12:30, Victoria Brennan <vkbrennan@hotmail.com> wrote:

> Dear Praveen,
>
> Please find attached a a completed PGR data collection fund application form. I have also attached the two related quotations. Please do let me know if you need any additional information from me. I look forward to hearing from you.
>
> Best Wishes,
>
> Victoria Brennan

Appendix 19: Process Utility Valuation Study

ScHARR Ethics Application

This form has been approved by the University Research Ethics Committee (UREC)

Date:	10th March 2013
Name of applicant:	Victoria Brennan
Research project title:	The addition of a process domain to the SF-6D

Complete this form if you are a **member of staff or a postgraduate research student** who plans to undertake a research project which requires ethics approval via the University Ethics Review Procedure. [X]

or

Complete this form if you plan to submit a **'generic' research ethics application (i.e. an application** that will cover several sufficiently similar research projects). Information on the 'generic' route is at: www.sheffield.ac.uk/ris/other/gov-ethics/ethicspolicy/approval-procedure/review-procedure/generic-research-projects [ ]

If you are an undergraduate or a postgraduate-taught student, this is the wrong form.

This form should be accompanied, where appropriate, by all Information Sheets/Covering Letters/Written Scripts which you propose to use to inform the prospective participants about the proposed research, and/or by a Consent Form where you need to use one.

Further guidance on how to apply is at: http://www.shef.ac.uk/scharr/research/ethicsgovernance

Guidance on the possible routes for obtaining ethics approval (i.e. on the University Ethics Review Procedure, the NHS procedure and the Social Care Research Ethics Committee, and the Alternative procedure) is at: www.shef.ac.uk/ris/other/gov-ethics/ethicspolicy/approval-procedure/ethics-approval

Once you have completed this research ethics application form in full, and other documents where appropriate, check that your name, the title of your research project and the date is contained in the footer of each page and email, as a word document, to the Ethics Administrator k.woodhead@sheffield.ac.uk.

Please note that the original signed and dated version of 'Part B' of the application form should be provided to the Ethics Administrator in hard copy.

I confirm that I have read the current version of the University of Sheffield 'Ethics Policy Governing Research Involving Human Participants, Personal Data and Human Tissue', as shown on the University's research ethics website at: www.shef.ac.uk/ris/other/gov-ethics/ethicspolicy

Yes

Part A

A1. Title of Research Project: Process Utility Survey

A2. Contact person (normally the Principal Investigator, in the case of staff-led research projects, or the student in the case of supervised-postgraduate researcher projects):

Title: Miss

Post: PGR student

Email: vkbrennan@hotmail.com

Name: Victoria Brennan

Department: HEDS, ScHARR

A3. Is this a postgraduate researcher project? If yes, please provide the Supervisor's contact details:

Title: Prof

Post: Prof Health Economics, HEDS

Email: s.dixon@sheffield.ac.uk

Name: Simon Dixon

Department: HEDS, ScHARR

Telephone: 0114 2220724

A3. Proposed Project Duration:

Start date: 25th April 2013 End date: 25th June 2013

A4. Mark 'X' in one or more of the following boxes if your research:

	involves adults with mental incapacity or mental illness
	involves prisoners or others in custodial care (e.g. young offenders)
	involves children or young people aged under 18 years
	involves using samples of human biological material collected before for another purpose
	involves taking new samples of human biological material (e.g. blood, tissue) *
	involves testing a medicinal product *
	involves taking new samples of human biological material (e.g. blood, tissue) *
	involves additional radiation above that required for clinical care *
	involves investigating a medical device *

	is social care research
	is ESRC funded
	Is taking place in the health service but does not require NHS ethical approval**
	URMS number if required (please see below)

* If you have marked boxes marked ***** then you also need to obtain confirmation that appropriate University insurance is in place. The procedure for doing so is entirely by email. Please send an email addressed to insurance@shef.ac.uk and request a copy of the 'Clinical Trial Insurance Application Form'.

- If you have marked the box** your supervisor, needs to obtain an URMS number (details on the ScHARR web site http://www.shef.ac.uk/scharr/research/ethicsgovernance/ugpgt)

It is recommended that you familiarise yourself with the University's Ethics Policy Governing Research Involving Human Participants, Personal Data and Human Tissue before completing the following questions. Please note that if you provide sufficient information about the research (what you intend to do, how it will be carried out and how you intend to minimise any risks), this will help the ethics reviewers to make an informed judgement quickly without having to ask for further details.

A5. Briefly summarise:

i. The project's aims and objectives:

(this must be in language comprehensible to a lay person)

Resources are limited and need to be allocated efficiently. In the United Kingdom (UK) this allocation process is carried out by the National Institute for Health and Clinical Excellence (NICE). NICE use a standard and internationally recognised method to compare different drugs and interventions and measure their clinical effectiveness. This is called the quality-adjusted life year (QALY) and is used as a common outcome measure to enable comparisons across different areas of health care.

In order to calculate QALYs utilities are required. These are a measure of preference for specific state of health quantified on a scale which is anchored by 0 representing "dead" and 1 representing "full health". Utility estimates are combined with the number of years spent in the health state to generate QALYs.

Several approaches can be used to estimate utility values for health states. One approach is the use of generic preference-based measures of health such as the SF-6D. The SF-6D is a standardised, multidimensional, generic measure of health related quality of life (HRQOL) with 6 domains (physical functioning, role limitations, social functioning, pain, mental health and vitality). Analysis of the SF-6D data provides utility estimates.

Patient's preferences for the processes involved in receiving care are an important issue in health care. For example, patients may prefer a less invasive test even if they must wait longer to receive it. However, no standardised questionnaire exists which captures this aspect of receiving care.

The study described here is a methodological piece of work with one main objective. To examine the feasibility of incorporating a process "bolt-on" question to the SF-6D. This would allow patients preferences for the processes of receiving care to be incorporated into utility estimates and so into the QALY calculation.

ii. The project's methodology:

(this must be in language comprehensible to a lay person)

Data will be collected using online surveys which will be administered to a representative sample of approximately 200 members of the UK general population. Two types of questions will be asked. Examples of the two question types can be seen in Figures 1 and 2. The system to be used for the delivery of health state valuation tasks has already been developed by staff at ScHARR in conjunction with a university spin off company for use in other projects.

An online participant panel will be used. Firstly, participants will read the project information and consent to take part. They will then complete twenty-one of the questions shown in Figure 1, followed by twenty-five of the questions shown in Figure 2. They will then be asked to complete questions on demographic background. The survey will take approximately 20 minutes to complete.

The cost for using the online participant panel, through Survey Sampling International is £762.00. This company has been used in previous studies undertaken at ScHARR. Funding has been obtained through the ScHARR PGR Data Collection Fund Application.

Figure 1.

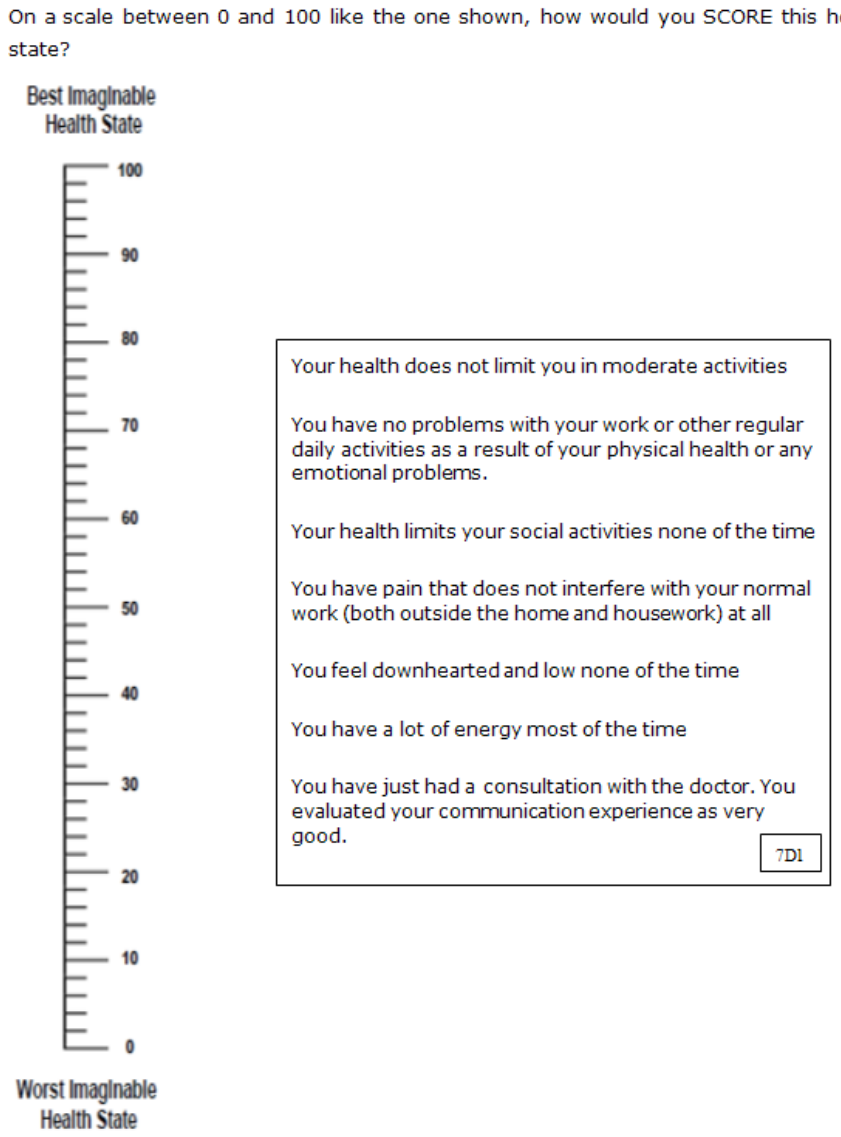
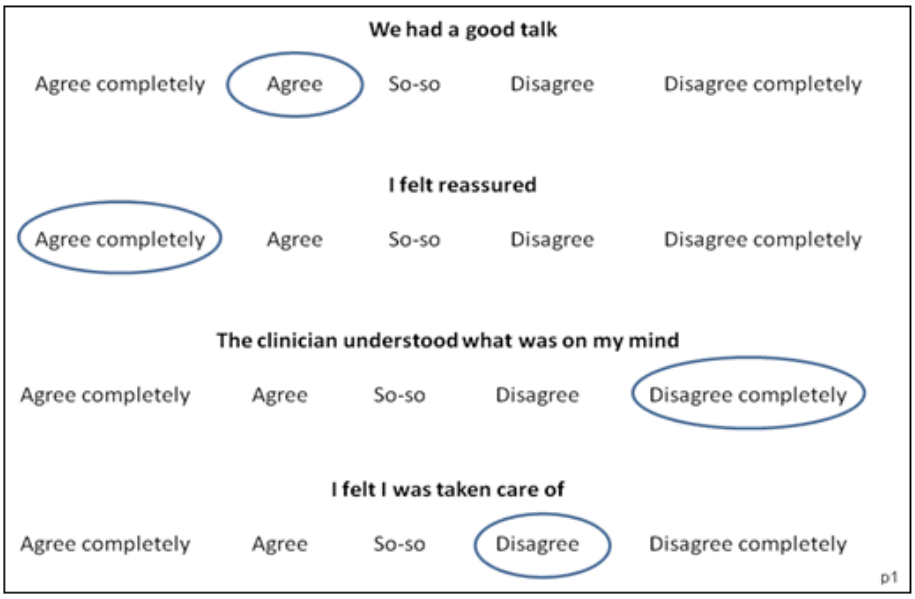


Figure 2



Would you consider this consultation to be?

Very poor ☐

Poor ☐

Average ☐

Good ☐

Very good ☐

The questions of the type shown in Figure 1 will generate values that represent how respondents rate the health state. Across the sample, we will then be able to estimate the degree to which these values change due to differences in their doctor consultation experience. The questions of the type shown in Figure 2 will allow us to statistically link responses to the Patient Experience Questionnaire (which supply the statements in Figure 2) to the overall assessment of the consultation. This will then allow us to apply the ratings of consultation experience to Patient Experience Questionnaire responses in a recently completed trial.

A6. What is the potential for physical and/or psychological harm/distress to participants?

There is minimal potential for harm during this project. A range of hypothetical health state scenarios that include a reference to death will be presented to participants. However, as we are asking participants to imagine themselves in the health state, we do not believe this has the potential for harm. Participants will be informed that they will answer questions relating to a hypothetical situation concerning their own death before taking part in the online survey or interview. Therefore, if this is something that concerns them, they will be able to make an informed decision over whether to take part.

A7. Does your research raise any issues of personal safety for you or other researchers involved in the project? (especially if taking place outside working hours or off University premises)

No

If yes, explain how these issues will be managed.

Not applicable

A8. How will the potential participants in the project be:

i. **Identified?** *(please ensure that all practical issues about contacting individuals are covered and that you are not requesting the personal details of individuals be given over without their consent)*

An online participant panel managed and maintained by an accredited market research

company will complete the survey. The sampling company has been used in previous studies undertaken by ScHARR. Participants will be randomly selected from those in the panel eligible to take part in this study who fulfil the age and gender quotas that are equivalent to the UK general population. All members of the panel have completed a double opt-in process to join the panel and have consented to complete online surveys over the course of their membership of the panel.

ii. Approached?

The participants randomly selected from the internet panel will be emailed with a link to take part in the survey. The email will include information about the survey so that participants can decide whether to participate in the survey or not.

iii. Recruited?

Participants will be recruited via an email containing information about the questionnaire and a link to the survey. A consent form is included in the online survey following the information pages, and there is no time limit for the completion of this page.

A9. Will informed consent be obtained from the participants?

Yes [X] No [ ]

If informed consent or consent is NOT to be obtained please explain why. Further guidance is at: www.shef.ac.uk/ris/other/gov-ethics/ethicspolicy/policy-notes/consent

Not applicable

A9.1. This question is only applicable if you are planning to obtain informed consent:

How do you plan to obtain informed consent? (i.e. the proposed process?):

After potential participants have read the study information and clicked the survey link, they will be presented with an online informed consent form that they will have to read and agree to, before proceeding to the questionnaire.

The Information Sheet details that participants are free to withdraw from the study if they wish. It also explains what would happen if they decide not to complete the survey.

Remember to attach your consent form and information sheet (where appropriate)

A10. What measures will be put in place to ensure confidentiality of personal data, where appropriate?

(As a minimum please ensure details are included of: how long data will be kept; when and how it will be destroyed; that PCs and other devices are password protected; that personal details are encrypted. This information should also be included on your information sheet).

The market research organisation that will provide the panel stores personal information such as email address and names in highly encrypted databases that are kept separate from the data collected as part of the study (with is also securely stored). Participants will only be identified using a unique ID number that cannot be linked to any person information stored by the panel provider. The project team will only use anonymised data collected as part of this study. The study data will be stored in an encrypted database and will also be kept on password protected university computers. Only aggregate level data will be reported in any report or publication arising from the project.

A11. Will financial/in kind payments (other than reasonable expenses and compensation for time) be offered to participants? (Indicate how much and on what basis this has been decided)

The panel provider supplying the participants for this study offer a small "thank you" for the completion of surveys as standard practice. This takes the form of points that are allocated dependent on the questionnaire length. Points can then be redeemed for a range of online vouchers. Using the points system online is seen as a neutral system that does not skew the participation of certain demographic groups.

A12. Will the research involve the production of recorded media such as audio and/or video recordings?

YES [ ] NO [X]

A12.1. This question is only applicable if you are planning to produce recorded media:

How will you ensure that there is a clear agreement with participants as to how these recorded media may be stored, used and (if appropriate) destroyed?

Not applicable.

Guidance on a range of ethical issues, including safety and well-being, consent and anonymity, confidentiality and data protection are available at: www.shef.ac.uk/ris/other/gov-ethics/ethicspolicy/policy-notes

University Research Ethics Application Form - Part B - The Signed Declaration

Title of Research Project:

Process Utility Survey

I confirm my responsibility to deliver the research project in accordance with the University of Sheffield's policies and procedures, which include the University's '*Financial Regulations*', '*Good Research Practice Standards*' and the '*Ethics Policy Governing Research Involving Human Participants, Personal Data and Human Tissue*' (Ethics Policy) and, where externally funded, with the terms and conditions of the research funder.

In signing this research ethics application form I am also confirming that:

- The form is accurate to the best of my knowledge and belief.

- The project will abide by the University's Ethics Policy.
- There is no potential material interest that may, or may appear to, impair the independence and objectivity of researchers conducting this project.
- Subject to the research being approved, I undertake to adhere to the project protocol without unagreed deviation and to comply with any conditions set out in the letter from the University ethics reviewers notifying me of this.
- I undertake to inform the ethics reviewers of significant changes to the protocol

  (by contacting my academic department's Ethics Administrator in the first instance).
- I am aware of my responsibility to be up to date and comply with the requirements of the law and relevant guidelines relating to security and confidentiality of personal data, including the need to register when necessary with the appropriate Data Protection Officer (within the University the Data Protection Officer is based in CiCS).
- I understand that the project, including research records and data, may be subject to inspection for audit purposes, if required in future.
- I understand that personal data about me as a researcher in this form will be held by those involved in the ethics review procedure (e.g. the Ethics Administrator and/or ethics reviewers) and that this will be managed according to Data Protection Act principles.
- If this is an application for a 'generic' project, all the individual projects that fit under the generic project are compatible with this application.
- **I understand that this project cannot be submitted for ethics approval in more than one department, and that if I wish to appeal against the decision made, this must be done through the original department.**
-

Name of the Supervisor:

Professor Simon Dixon

Name of the student:

Victoria Brennan

Signature of the Supervisor:

[redacted]	**Date: 4th April 2013**

Email the completed application form and provide a signed, hard copy of 'Part B' to the Ethics Administrator (also enclose, if relevant, other documents).

Appendix 20: Process Utility Valuation Study Protocol

The addition of a process domain to the SF-6D

Study Protocol.

March 2013

Prepared by Victoria Brennan

Introduction

Resources are limited and need to be allocated efficiently. In the United Kingdom (UK) this allocation process is carried out by the National Institute for Health and Clinical Excellence (NICE). NICE use a standard and internationally recognised method to compare different drugs and interventions and measure their clinical effectiveness. This is called the quality-adjusted life year (QALY) and is used as a common outcome measure to enable comparisons across different areas of health care (Drummond et al., 2005).

In order to calculate QALYs utilities are required. These are a measure of preference for specific state of health quantified on a scale which is anchored by 0 representing "dead" and 1 representing "full health". Utility estimates are combined with the number of years spent in the health state to generate QALYs (Drummond et al., 2005).

Several approaches can be used to estimate utility values for health states. One approach is the use of generic preference-based measures of health. Greater utilisation of these generic measures has highlighted that for some diseases they fail to capture important dimensions of health. As a result of this, several health economists have attempted to add or "bolt-on" dimensions to existing instruments. Studies have been performed in 3 main treatment areas: cognition, sleep and pain. These studies will now be discussed in greater depth (Krabbe et al., 1999; Wolfs et al., 2007; Yang et al., 2008; Brazier et al., 2010).

The context of this study is a trial of an electronic patient assessment questionnaire (ePAQ) for use in the care of women with pelvic floor problems. The benefits of ePAQ are thought to be restricted to patient satisfaction with the clinical assessment process. Such benefits are not expected to be captured by a generic preference-based measure, and so a "bolt-on" that captures this potential source of process utility has been developed and will be valued in this study.

Bolt-on Studies

Krabbe et al. (1999)

Krabbe et al (1999) undertook a study to add a cognitive dimension to the EQ-5D-3L. The EQ-5D contains 5 dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). Krabbe et al (1999) question the relevance of the EQ-5D for use in studies of dementia patients considering the absence of a cognition domain. They assumed the addition of a cognitive dimension would increase the tool's comprehensiveness.

They maintained the standard EQ-5D-3L classification system, and added a 6th cognition domain also with the 3 levels (no impairment of cognitive functioning, some impairment of cognitive functioning and severe impairment of cognitive functioning). The researchers selected 14 health states ensuring a representative coverage of the scale, and added the sixth attribute. The best and worst health states were also included. Two additional EQ-5D states were added to construct a parallel set of EQ-5D+C health states with an assumed significant effect of the cognitive functioning attribute. The researchers added two additional health states to the EQ-5D set which differed over only 1 level in cognition from an already selected EQ-5D+C state. The resulting study contained 18 health states for the EQ-5D and 20 health states for the EQ-5D+C.

Health state descriptions were valued using the visual analogue scale (VAS) which uses a rating scale from 0 to 100, where 0 is the "worst imaginable health state" and 100 is the "best imaginable health state". Respondents were asked to locate the health state descriptions on the scale between the two anchors so that the intervals between the scores corresponded to the differences they perceive.

The distribution of EQ-5D questionnaires and EQ-5D+C questionnaires was performed randomly. 185 questionnaires were sent out. 87 valid questionnaires were returned: 39 having completed the EQ-5D version and 48 having completed the EQ-5D+C version.

Analysis of the data included: Calculations of means and standard deviations for all valued health-state descriptions; Statistical testing within pairs (EQ-5D vs EQ-5D+C); Paired t-tests were performed for some health states.

Results of t-tests for differences between the EQ-5D and EQ-5D+C showed significant differences in 2 cases: where a "moderate" cognition level was added to a "good" state, and where a "bad" cognition level was added to a "moderate" state. Adding a "good" cognition level to a "bad" health state did not result in any meaningful improvements. Adding a "bad" cognition level to a "bad" health state resulted in a poor valuation. Therefore indicating that "good" health states were affected more by the addition of impaired cognition, but the reverse was not true.

Authors included some health states twice in the study [((11111(1)), (33333(3))] and compared valuations for both pairs within each version as a reliability test. Paired t tests showed no difference in valuations of 11111(1) for either version but a small but significant difference in valuation for health state 333333 for the EQ-5D+C version.

The authors concluded that the approach they developed to add the cognition domain to the existing EQ-5D was suitable for judging the revised EQ-5D+C tool, and that those who are developing multi-attribute health status tools in the future should be aware of the omission of cognition in the existing tools available.

Wolfs et al. (2007)

Wolfs et al. (2007) used both the EQ-5D-3L and EQ-5D-3L+C in a population of patients with cognitive impairment, and compared the validity of both tools through assessment of construct validity. The EQ-5D and VAS were administered to patient proxies, who were then asked to value the additional cognitive domain, and complete the VAS. Patient proxies also completed the MMSE (mini mental state examination). The study concluded that the construct validity of the EQ-5D and EQ-5D+C were comparable except for the VAS-5D. Correlations between the cognitive dimension of the MMSE were most similar to the correlations between the self-care and usual activities dimension and the MMSE. The EQ-5D was found to be slightly more responsive than the EQ-5D+C. The authors concluded that the EQ-5D performed well for evaluating HRQOL in patients with cognitive impairment, and that the addition of the cognition domain was not necessary. The authors did however, discuss the limitations to their study, namely, that the exploration into the two tools was performed within a RCT which was not designed purely for this purpose, and therefore improvements could have been made to the methodology

i.e. patient proxies complete the EQ-5D and the EQ-5D+C (not just the additional domain) and the use of patient proxies, rather than the patients themselves.

Yang et al. (2008)

Yang et al (2008) undertook a study examining the effect of adding a dimension sleep dimension to the EQ-5D. The additional dimension maintained the 3 levels within the EQ-5D (I have no problems with sleep; I have some problems with sleep; I have extreme problems with sleep).

Sixteen EQ-5D and 18 EQ-5D+sleep health states were developed, and within these sets of states there were 6 matched health state pairs (a matched pair contained an EQ-5D state and a corresponding EQ-5D+sleep state which only differs with the addition of the sleep dimension). The EQ-5D and EQ-5D+sleep health states were then valued by the general population using TTO techniques. This study did not include the VAS. Comparisons between paired states suggested that adding the sleep dimension had no influence on valuations of the original EQ-5D health state.

Brazier et al. (2010)

Brazier et al., (2010) explored the impact of adding a pain dimension to the AQL-5D. The AQL-5D is an asthma specific preference based measure derived from the Asthma Quality of Life questionnaire (AQLQ). The original AQLQ contains 5 domains (concern about asthma, shortness of breath, weather and pollution stimuli, sleep impact and activity limitations), and each have 5 levels (1 = no problem, 5 = severe problems). When mapped onto the AQLD-5D this results in a health state classification system with 3 level of severity (1 = no problems, 2 = some problems, 3 = extreme problems), and this reduced health-state classification system was used in the study. Brazier et al (2010) added the pain domain from the EQ-5D to the classification system to create the AQL-6D. They then performed a valuation study of both the AQL-5D and the AQL-6D. 16 health states were selected for the

AQL-5D and 18 health states were selected for the AQL-6D. These included 4 health states that were matched across each measure. Members of the public valued 8 health states from either the AQL-5D or AQL-6D using TTO. The impact of adding the dimension was then examined by comparing the mean values for the matched states, and by modelling the data. The results showed that the additional pain dimension had a significant coefficient and impacted significantly on the coefficients of the other dimensions. However, the degree of impact differed by dimension and severity level. Brazier et al (2010) applied their findings to clinical trial data. Values generated for the AQL-5D and AQL-6D were compared in terms of mean scores for all patients included in the trial, and also for sub-samples of patients identified according to asthma symptom scores. The results found that the mean health state value produced from the ALQ-5D was consistently lower than for the AQL-6D value across the 5 asthma symptom severity groups. For the pain dimension, the AQL-6D group mean scores only exceeded the AQL-5D scores for extreme pain.

The proposed methods for integrating process utility into the utility function draw on these previously performed studies, and attempt to overcome some of the key issues they discuss. The study furthers the common thesis theme of improving the relevance of trial outcomes to enhance the cost-effectiveness outcomes.

Background

This study aims to utilise SF-6D, Patient Experience Questionnaire (PEQ) and Client Satisfaction Questionnaire (CSQ-8) data collected within the ePAQ trial, to investigate the presence, and magnitude of process utility. The use of these data draws on 3 primary advantages:

- The SF-6D, PEQ and CSQ-8 are established instruments, which have been validated using psychometric testing, thereby testing for reliability and validity.
- The avoidance of using the valuation of hypothetical health states aims to reduce the focusing effect reported within the Brazier (2010) study.

- The increase in internal validity (how so we know that the differences identified are a true representation of what actually happened). The "true" causes of the outcomes that were observed in the ePAQ trial.

The method of identifying and measuring process utility has 4 key stages:

Stage 1. Defining process utility: The identification of the most appropriate aspect(s) of the PEQ and/or CSQ to use to represent "process" and incorporate into the SF-6D;

Stage 2. The classification and description the SF-6D plus process utility (SF-7D) health states;

Stage 3. Valuation study to empirically compare the SF-6D and SF-7D;

Stage 4. Application in practice: relating the findings back to the ePAQ RCT.

Stages 1 and 2 have been completed as preparatory work to the health state valuation study (stage 3) and application in practice (stage 4).

Stage 1: Defining the Process Utility Domain.

Exploratory Analysis

Initially, exploratory analysis of ePAQ trial data was performed. A Mann-Whitney test was performed to test for differences between PEQ and CSQ-8 responses for the 2 trial arms. The presence of a significant difference between arms was assumed to reflect a difference induced by the process of care (as a key difference between the trial arms was the mode of consultation: face-to-face versus telephone). The results showed a significant difference between 3 PEQ questions and 2 PEQ domains. There was no significant difference between any of the CSQ-8 items. Further analysis therefore concentrated only on the PEQ.

Pilot Study

Following the exploratory analysis, a pilot study was performed. The aim of the pilot study was to determine how respondents prefer to have the communication domain portrayed to them, and to determine whether it was possible to ask respondents to rate the PEQ health state, and calculate a rating score for each combination of PEQ responses.

All individuals working for RTI-Health Solutions and based in the UK offices were sent interview booklets and asked if they would like to participate in the pilot study. The interview booklets contained 3 components:

- Respondents were shown 3 SF-7D health states . The process domain was represented by one question taken directly from the PEQ questionnaire. They were asked to value these using the VAS.
- Respondents were shown 3 further SF-7D health states, this time represented as a rating score, (Methodology B: Appendix A: Figure A-2), and asked to value the health states using the VAS.
- Respondents were asked which set of descriptions they preferred.
- Respondents were shown 4 scenarios' which were based on the communication domain of the PEQ (Appendix A: Figure A-3). They were informed that the scenario's described their feelings immediately after a consultation with the doctor. For each scenario they were asked if they would consider the consultation to be very poor, poor, fair, good or very good.

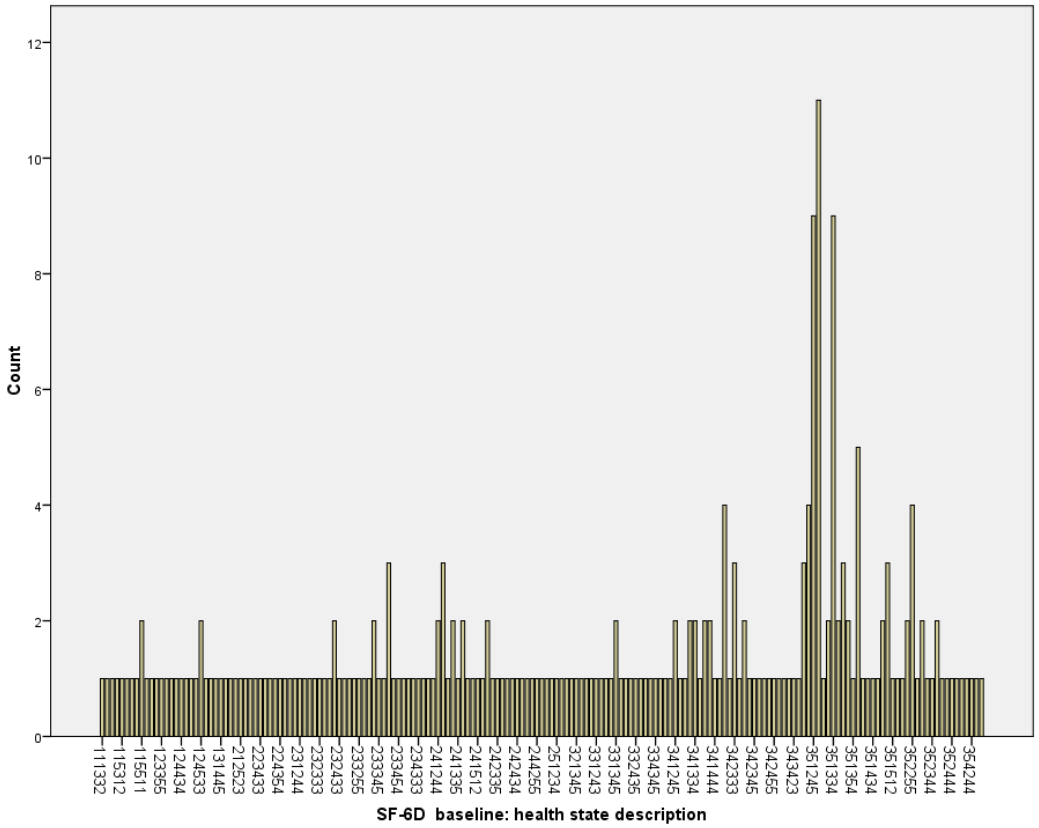
Twenty-three individuals returned completed questionnaires. Of these 8 preferred methodology A and 15 preferred methodology B. This introduced the requirement that responses to the PEQ communication domain questions need to be converted to an overall rating for the domain. If appropriate data are collected, multinomial regression can be used to predict the rating for the overall communication experience based on responses to the individual items within the communication domain. A fractional factorial design is used to identify a set of communication item combinations that will allow overall communication experience to be estimated statistically.

Stage 2: The classification and description the SF-6D plus process utility (SF-7D) health states

SF-6D data collected from patients within the ePAQ trial were used to determine the health states to be valued. The results from the valuation study will be related back to the ePAQ trial data, therefore the initial selection criteria for health states is that they must be those experienced by trial patients. The SF-6D health states experienced by patients and the frequency by which they were experienced are represented in 0. There were 252 health states experienced by patients within the trial. Although it is not possible to identify every health state within the histogram it is evident that the majority of patients experienced the less severe SF-6D health states.

The second selection criterion mirrors Krabbe et al., (1999) who selected states purely to ensure an even mix of good, moderate and bad health states based on the EQ-5D value. A total of 20 health states were selected. The health states experienced ranged from 111112 (utility estimate of 0.922) to 34555 (utility estimate of 0.345). Therefore, health states included were also within this range. The most commonly reported health state was 111112 (n=11). The seventh attribute was added to all of these 20 health states, again, selection was based on methodology used by Krabbe et al., (1999) whereby levels were chosen randomly, occasionally avoiding a level too unlikely in combination with the other six levels. A further two SF-7D health states were added differing only one level in the process domain from an already selected SF-7D state.

Figure 35. Histogram of SF-6D Health States Experienced in ePAQ Trial.



The remainder of this study protocol refers to stage 3: The valuation study to empirically compare the SF-6D and SF-7D, and presents the proposed methodology and analysis plan.

Objective

The objective of this health state valuation study is to examine the effect of extending the SF-6D to include a seventh "process" domain.

Methodology

The proposed methodology is largely based on the approach taken by Krabbe et al., (1999), however, it furthers this approach through the utilization of data already collected as part of the ePAQ trial.

As detailed in Section 2, two sets of health states have been developed:

- SF-6D health states;
- SF-7D health states (these mirror the SF-6D health states, and also include a seventh process domain).

Valuations of both sets of health states using the VAS will be analysed to detect the effect of the additional dimension.

Survey

A health state valuation survey will be administered to a sample of the UK general population. The system used for the delivery of the health state valuation tasks has already been developed by staff at ScHARR in conjunction with a university spin off company for use in other projects. An online participant panel will be used to source respondents.[1]

Firstly participants will be asked to read the project information and consent to take part. A copy of the project information and patient consent form can be seen in Appendix D.

After providing consent to participate in the study, respondents will be asked to complete a survey. The survey will take approximately 25 minutes to complete. Two versions of the survey will be administered ("Block 1" and "Block 2") at random to the sample . A paper version of the survey – containing both Block 1 and 2 questions - can be seen in Appendix E. The survey includes three parts:

1. SF-6D(7D) health states.

[1] The cost for using the online participant panel, through Survey Sampling International is £762.00. This company has been used in previous studies undertaken at ScHARR. Funding has been obtained through the ScHARR PGR Data Collection Fund Application.

Respondents will be shown either a set of 21 SF-6D health states ("Block 1"), or a set of 23 SF-7D health states ("Block 2") and be asked to value them using the VAS.

2. PEQ health states with ratings.

   Respondents will be shown a set of 25 scenarios describing their feelings immediately after a consultation with the doctor (defined using the communication domain of the PEQ and the health states identified using the fractional factorial design detailed in section 2.1). Respondents will be asked whether they would consider the communication experience to be very good, good, fair, poor or very poor.

3. Demographic details.

   The final part of the survey includes demographic questions including age, gender, ethnicity, marital status and education.

Study Population.

The survey will be administered to 200 members of the UK general population. This sample size is based on Krabbe et al., (1999) who sent out 185 paper questionnaires and achieved a 47% response rate, resulting in 87 valid questionnaires being returned. They discuss their sample size as a key limitation to their study.

The sample size of 200 was based on Krabbe et al., (1999), discussions with my supervisor and budgetary constraints.

An online participant panel managed and maintained by an accredited market research company will complete the survey. Participants will be randomly selected from those in the panel eligible to take part in this study who fulfil the age and gender quotas that are equivalent to the UK general population. All members of the panel have completed a double opt-in process to join the panel and have consented to complete online surveys over the course of their membership of the panel.

The participants randomly selected from the internet panel will be emailed with a link to take part in the survey. The email will include information about the survey so that participants can decide whether to participate in the survey or not.

Data analysis

All analyses will be performed using SPSS. The following analysis will be presented:

- Demographic characteristics of the two groups of participants (those completing the SF-6D and those completing the SF-7D) will be summarised.
- Analysis of the SF-6D and SF-7D health state valuations
    - Means and standard deviations for all valued SF-6D and SF-7D health state descriptions will be reported.
    - Statistical testing within pairs (SF-6D versus SF-7D) of the descriptions will be performed using t tests and regressions to take account of respondent characteristics.
    - Statistical testing of health states that differ just in their communications levels using paired t tests and regressions to take account of respondent characteristics.
    - These analyses will allow us to determine the differences in utility associated with the different ratings of the communication experience.
- Analysis of the PEQ health state and rating scores
    - The multinomial regression will be undertaken using data collected from within the survey to predict overall communication experience based on individual PEQ item responses.
    - Results of the multinomial regression will be applied to the ePAQ trial data to obtain a rating of the communication experienced within the ePAQ trial consultation.
    - For those ePAQ trial patients who have completed the SF-6D, utility values can then be estimated using both the standard SF-6D tariff and incorporating differences relating to the process domain.

References

Brazier, J., Rowen, D., Tsuchiya, A., Yang, Y., & Young, T. 2010. What a pain: adding a generic dimension to a condition-specific preference-based measure. HESG Abstract

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes. Third Edition. Oxford University Press, UK; 2005.

Field, A. 2009. Discovering Statistics using SPSS, Third ed. USA, Sage Publishing.

Hicks, C. 2004. Research Methods for Clinical Therapists, Fourth ed. London, Churchill Livingstone.

Krabbe, P., Stouthard, M., Essink-Bot, M., & Bonsel, G. 1999. The effect of adding a cognitive dimension to the EuroQol Multiattribute Health-Status Classification System. J Clin Epidemiol, 52, (4) 293-301

Wolfs, C., Dirksen, C., Kessels, A., Willems, D., Verhey, F., & Severerns, J. 2007. Performance of the EQ-5D and the ED-5D+C in elderly patients with cognitive impairments. Health and Quality of Life Outcomes, 5, 33

Yang, Y., Brazier, J., & Tsuchiya, A. 2008. The effect of adding a 'sleep' dimension to EQ-5D. HESG

Appendix 21: Summary of Selected Health States

Table 40. **Selected SF-6D (7D) Health States.**

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
1	111112(1)	0.922	11	• Your health does not limit you in moderate activities • You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems. • Your health limits your social activities none of the time • You have pain that does not interfere with your normal work (both outside the home and housework) at all • You feel downhearted and low none of the time • You have a lot of energy most of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as very good).
1a	111112(2)	0.922		• Your health does not limit you in moderate activities • You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems. • Your health limits your social activities none of the time • You have pain that does not interfere with your normal work [both outside the home and housework] at all

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
				• You feel downhearted and low none of the time • You have a lot of energy most of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as good).
2	111113(1)	0.922	5	• Your health does not limit you in moderate activities • You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems. • Your health limits your social activities none of the time • You have pain that does not interfere with your normal work [both outside the home and housework] at all • You feel downhearted and low none of the time • (You have a lot of energy none of the time You have just had a consultation with the doctor. You evaluated your communication experience as very good).
3	111212(1)	0.922	4	• Your health does not limit you in moderate activities • You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems. • Your health limits your social activities none of the time • You have pain that interferes with your normal work [both outside the home and housework] a little bit • You feel downhearted and low none of the time • You have a lot of energy most of the time

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
				• (You have just had a consultation with the doctor. You evaluated your communication experience as very good).
4	111122(3)	0.863	6	• Your health does not limit you in moderate activities • You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems. • Your health limits your social activities none of the time • You have pain that does not interfere with your normal work [both outside the home and housework] at all • You feel downhearted and low some of the time • You have a lot of energy a little of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as average).
5	221112(3)	0.859	1	• Your health limits you a little in moderate activities • You are limited in the kind of work or other activities as a result of your physical health • Your health limits your social activities none of the time • You have pain that does not interfere with your normal work [both outside the home and housework] at all • You feel downhearted and low none of the time • You have a lot of energy most of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as average).

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
6	121122(3)	0.8	3	• Your health does not limit you in moderate activities • You are limited in the kind of work or other activities as a result of your physical health • Your health limits your social activities none of the time • You have pain that does not interfere with your normal work [both outside the home and housework] t all • You feel downhearted and low a little of the time • You have a lot of energy most of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as average).
6a	121122(4)	0.8	3	• Your health does not limit you in moderate activities • You are limited in the kind of work or other activities as a result of your physical health • Your health limits your social activities none of the time • You have pain that does not interfere with your normal work [both outside the home and housework] at all • You feel downhearted and low a little of the time • You have a lot of energy most of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as poor).

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
7	221122(2)	0.8	3	• Your health limits you a little in moderate activities • You are limited in the kind of work or other activities as a result of your physical health • Your health limits your social activities none of the time • You have pain that does not interfere with your normal work [both outside the home and housework] at all • You feel downhearted and low a little of the time • You have a lot of energy most of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as good).
8	241212(5)	0.782	1	• Your health limits you a little in moderate activities • You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems • Your health limits your social activities none of the time • You have pain that interferes with your normal work [both outside the home and housework] at a little bit. • You feel downhearted and low none of the time • You have a lot of energy most of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as very poor).

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
9	241224(1)	0.723	2	• Your health limits you in moderate activities • You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems. • Your health limits your social activities none of the time • You have pain that interferes with your normal work [both outside the home and housework] at a little bit. • You feel downhearted and low a little of the time • You have a lot of energy a little of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as very good).
10	131132(4)	0.722	4	• Your health does not limit you in moderate activities • You accomplish less than you would like as a result of emotional problems. • Your health limits your social activities none of the time • You have pain that does not interfere with your normal work [both outside the home and housework] at all • You feel downhearted and low some of the time • You have a lot of energy most of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as poor).
11	343112(3)	0.671	1	• Your health limits you alot in moderate activities

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
				• You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems. • Your health limits your social activities some of the time • You have pain that does not interfere with your normal work [both outside the home and housework] at all. • You feel downhearted and low none of the time • You have a lot of energy most of the time. • (You have just had a consultation with the doctor. You evaluated your communication experience as average).
12	132133(4)	0.66	4	• Your health does not limit you in moderate activities • You accomplish less than you would like as a result of emotional problems. • Your health limits your social activities a little of the time • You have pain that does not interfere with your normal work [both outside the home and housework] at all • You feel downhearted and low some of the time • You have a lot of energy some of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as poor).
13	143233(5)	0.657	3	• Your health does not limit you in moderate activities • You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emo-

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
				tional problems. • Your health limits your social activities some of the time • You have pain that interferes with your normal work [both outside the home and housework] a little bit. • You feel downhearted and low some of the time • You have a lot of energy some of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as very poor).
14	243223(2)	0.657	1	• Your health limits you a little in moderate activities. • You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems. • Your health limits your social activities some of the time • You have pain that interferes with your normal work [both outside the home and housework] a little bit. • You feel downhearted and low a little of the time • You have a lot of energy some of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as good).
15	243334(2)	0.615	4	• Your health limits you a little in moderate activities. • You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emo-

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
				tional problems. • Your health limits your social activities some of the time • You have pain that interferes with your normal work [both outside the home and housework] moderately. • You feel downhearted and low some of the time • You have a lot of energy none of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as good).
16	243434(5)	0.58	1	• Your health limits you a little in moderate activities. • You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems. • Your health limits your social activities some of the time • You have pain that interferes with your normal work [both outside the home and housework] quite a bit. • You feel downhearted and low some of the time • You have a lot of energy a little of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as very poor).
17	343423(2)	0.535	1	• Your health limits you a lot in moderate activities

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
				• You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems. • Your health limits your social activities some of the time • You have pain that interferes with your normal work [both outside the home and housework] quite a bit. • You feel down hearted and low a little of the time • You have a lot of energy some of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as good).
18	343435(5)	0.507	2	• Your health limits you a lot in moderate activities • You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems. • Your health limits your social activities some of the time • You have pain that interferes with your normal work [both outside the home and housework] quite a bit. • You feel downhearted and low some of the time • You have a lot of energy none of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as very poor).

#	Health State SF-6D(7D)	SF-6D score	Frequency of SF-6D in ePAQ trial	Health State Description SF-6d (7D)
19	345535(4)	0.42	1	• Your health limits you a lot in moderate activities • You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems. • Your health limits your social activities all of the time • You have pain that interferes with your normal work [both outside the home and housework] extremely • You feel downhearted and low some of the time • You have a lot of energy none of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as poor).
20	345555(5)	0.345	2	• Your health limits you a lot in moderate activities • You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems. • Your health limits your social activities all of the time • You have pain that interferes with your normal work [both outside the home and housework] extremely • You feel downhearted and low none of the time • You have a lot of energy none of the time • (You have just had a consultation with the doctor. You evaluated your communication experience as very poor).

Appendix 22: Patient information and consent form.

Introduction Page

Thank you for your interest in participating in the Process Utility survey. Before you decide whether to complete this survey it is important that you understand why the survey is being carried out and what it will involve. Please read the information provided carefully and if you are willing to take part please click the continue button below.

Why is the survey being carried out?

Different people experience different states of health. Some people's health is better than others. We are interested in comparing different states of health and measuring how good or bad they are.

This questionnaire is part of a study to determine the influence of the process of receiving health care has on a person's quality of life.

Why have I been chosen to take part?

You have been randomly selected by Survey Sampling International as you are a member of one of their online panels and are eligible to take part in the project. There will be 200 people taking part.

Do I have to take part?

It is up to you to decide whether or not to take part. If you do not want to take part it will not affect you in any way. If you agree to take part you will need to complete an online informed consent form before accessing the survey. This means that you have agreed to complete the survey. If you agree to take part but then decide that you do not want to finish the survey you are free to stop at any time.If you decide to stop the survey at any time, the study team may still use your responses to any question you have answered. If you wish to withdraw from the study altogether, then you can contact the study team at vbrennan@rti.org. Once you do this, your responses will not be used in analysis.

What will happen to me if I take part?

If you take part in the survey you will be asked a number of questions. The survey will last about 20-25 minutes. This will consist of three parts. In the first part you will be asked to think about a set of 21 imaginary health states and tell us how good or bad you think they are. Some of the questions may be of a sensitive nature, and may ask you to think about difficult issues, such as death or dying.

In the second part we will then ask you to think about 25 scenarios describing your feelings immediately after a consultation with the doctor, and tell us if you would consider the communication experience to be very good, good, fair, poor or very poor.

The third part of the survey asks you some questions about you and your health. All of your responses will be treated as confidential, and all analysis will be carried out anonymously. We are interested in people's views, and there are no right or wrong answers. Please tell us what you think. Upon successful completion of the survey you will be returned to your panel platform and if eligible receive points for completing this survey as per usual.

If you find some of the questions upsetting, or if you wish to seek advice or reassurance about your own health, then either contact your GP or NHS Direct (Tel: 0845 46 47).

Will my taking part in this project be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential. The information you give will not be used in any way that could identify you.

What will happen to the results of the research project?

The results of this study will be published in academic journals and presented at conferences. Nobody will be able to identify you in any reports or publications as only aggregate results will be published. If you would like a copy of the results once the study has finished please contact the research team led by Professor Simon Dixon, via Ann Hilton (a.hilton@shef.ac.uk).

Who is organising and funding the research?

Survey Sampling International are organising the online survey on behalf of the University of Sheffield. The project has been funded by the University of Sheffield.

Who has ethically reviewed the project?

This research project has been reviewed by the School of Health and Related Research Ethics Committee at the University of Sheffield.

If you are having technical difficulties accessing the survey please click here

If you have any questions about the research, please contact:

Victoria Brennan (xx)

Participant Consent Form

Title of Project: Process Utility Survey

Name of Researchers: Victoria Brennan, University of Sheffield

Respondent ID for this project: AUTOMATIC

Please tick box

1. I confirm I have read + understand the project information provided + if I require further details I can contact the researchers by email ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without having to give a reason and without any sort of penalty ☐

3. I understand that any data I provide will be treated securely and kept confidential. My responses will be anonymised before analysis. Only the project team will have access to my responses. ☐

4. I agree to take part in the Process Utility Survey. ☐

5. I am a resident of the United Kingdom and I am over 18 years of age ☐

If you would like any further information about the research please e-mail either Professor Simon Dixon (s.dixon@shef.ac.uk) or Victoria Brennan (v.brennan@shef.ac.uk).

To continue to the survey please click continue below.

Appendix 23: ScHARR Ethics Committee Approval, Process Utility Valuation.

The University Of Sheffield.

Kirsty Woodhead
Ethics Committee Administrator

Regent Court
30 Regent Street
Sheffield S1 4DA
Telephone: +44 (0) 114 2225453
Fax: +44 (0) 114 272 4095 (non confidential)
Email: k.woodhead@sheffield.ac.uk

Our ref: 0642/KW

22 May 2013

Victoria Brennan
ScHARR

Dear Victoria

Process Utility Brennan

Thank you for submitting the above research project for approval by the ScHARR Research Ethics Committee. On behalf of the University Chair of Ethics who reviewed your project, I am pleased to inform you that on 22 May 2013 the project was approved on ethics grounds, on the basis that you will adhere to the documents that you submitted for ethics review.

The research must be conducted within the requirements of the hosting/employing organisation or the organisation where the research is being undertaken. You are also required to ensure that you meet any research ethics and governance requirements in the country in which you are researching. It is your responsibility to find out what these are.

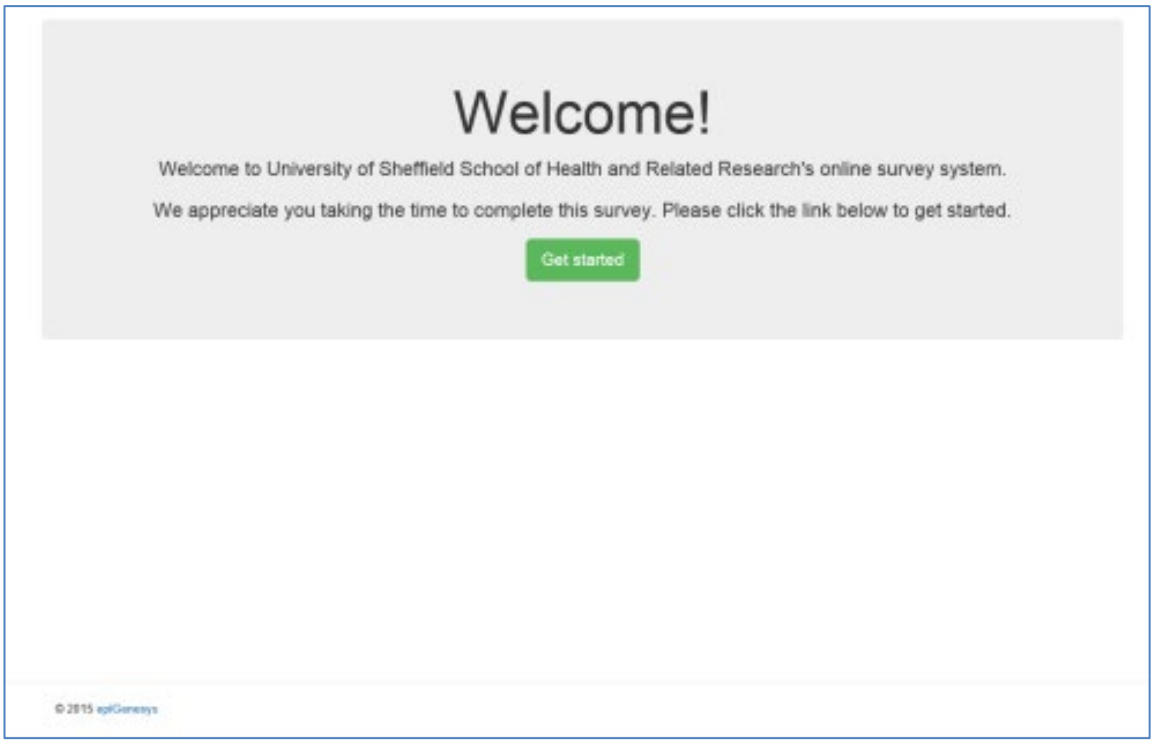
If during the course of the project you need to deviate significantly from the documents you submitted for review, please inform me since written approval will be required. Please also inform me should you decide to terminate the project prematurely.

Yours sincerely

[redacted]

Kirsty Woodhead
Ethics Committee Administrator

Appendix 24: Process Utility Online Survey.



ScHARR Survey

Process Utility survey

Thank you for your interest in participating in the Process Utility survey. Before you decide whether to complete this survey it is important that you understand why the survey is being carried out and what it will involve. Please read the information provided carefully and if you are willing to take part please click the continue button below.

Why is the survey being carried out?

Different people experience different states of health. Some people's health is better than others. We are interested in comparing different states of health and measuring how good or bad they are.

This questionnaire is part of a study to determine the influence that the way patients receive health care has on a person's quality of life.

Why have I been chosen to take part?

You have been randomly selected by Survey Sampling International as you are a member of one of their online panels and are eligible to take part in the project. There will be 200 people taking part.

Do I have to take part?

It is up to you to decide whether or not to take part. If you do not want to take part it will not affect you in any way. If you agree to take part you will need to complete an online informed consent form before accessing the survey. This means that you have agreed to complete the survey. If you agree to take part but then decide that you do not want to finish the survey you are free to stop at any time.

If you decide to stop the survey at any time, the study team may still use your responses to any question you have answered. If you wish to withdraw from the study altogether, then you can contact the study team on processsurvey@sheffield.ac.uk. Once you do this, your responses will not be used in analysis.

ScHARR Survey

What will happen to me if I take part?

If you take part in the survey you will be asked a number of questions. The survey will last about 20-25 minutes. This will consist of three parts. In the first part you will be asked to think about a set of 21 or 23 imaginary health states and tell us how good or bad you think they are. Some of the questions may be of a sensitive nature, and may ask you to think about difficult issues, such as death or dying.

In the second part we will then ask you to think about 20 scenarios describing your feelings immediately after a consultation with the doctor, and tell us if you would consider the communication experience to be very good, good, fair, poor or very poor.

The third part of the survey asks you some questions about you and your health. All of your responses will be treated as confidential, and all analysis will be carried out anonymously. We are interested in people's views, and there are no right or wrong answers. Please tell us what you think. Upon successful completion of the survey you will be returned to your panel platform and if eligible receive points for completing this survey as per usual.

If you find some of the questions upsetting, or if you wish to seek advice or reassurance about your own health, then either contact your GP or NHS Direct (Tel: 0845 46 47). Please note that we do not expect any participants to be distressed.

Will my taking part in this project be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential. The information you give will not be used in any way that could identify you.

ScHARR Survey

What will happen to the results of the research project?

The results of this study will be published in academic journals and presented at conferences. Nobody will be able to identify you in any reports or publications as only aggregate results will be published. If you would like a copy of the results once the study has finished please contact the research team led by Professor Simon Dixon, via Ann Hilton (a.hilton@shef.ac.uk).

Who is organising and funding the research?

Survey Sampling International (SSI) are organising the online survey on behalf of the University of Sheffield. The project has been funded by the University of Sheffield.

Who has ethically reviewed the project?

This research project has been reviewed by the School of Health and Related Research Ethics Committee at the University of Sheffield.

If you have any questions about the research, or any issues or complaints related to the survey, please contact the research team on:

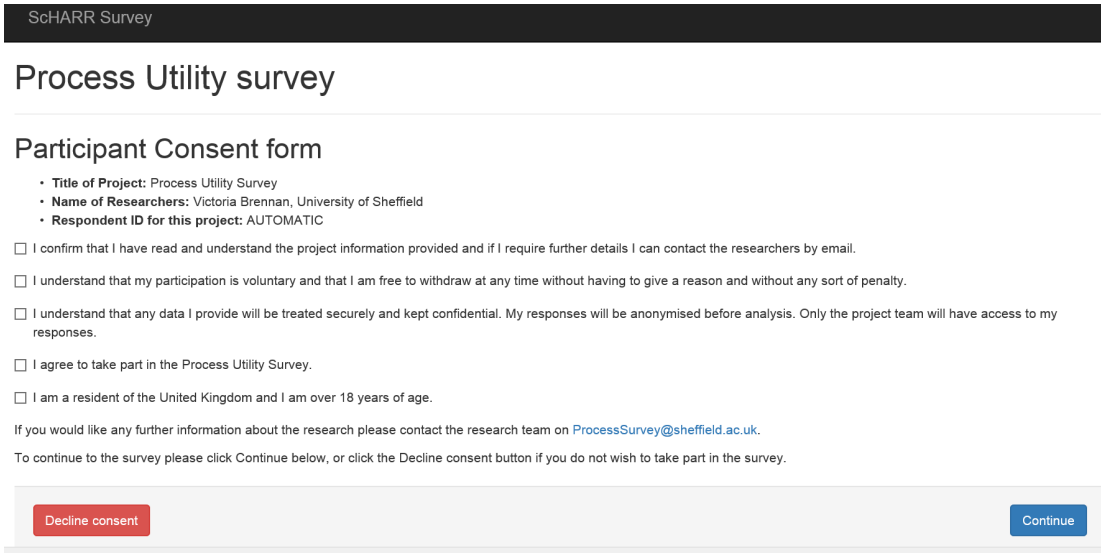
ProcessSurvey@sheffield.ac.uk.

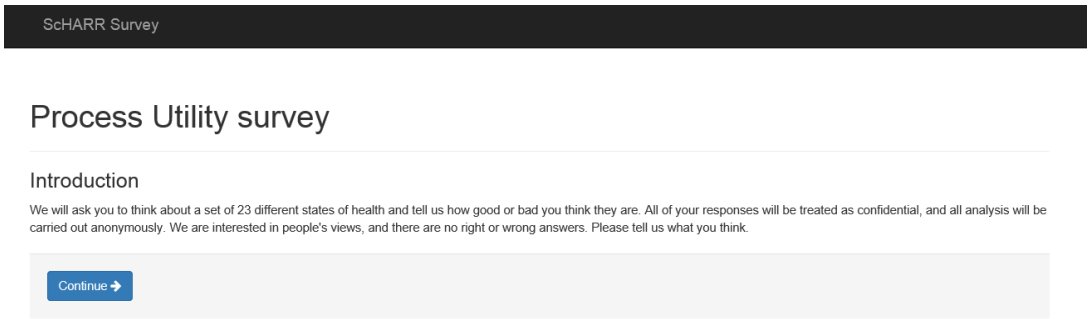
If you are having technical difficulties accessing the survey please contact:

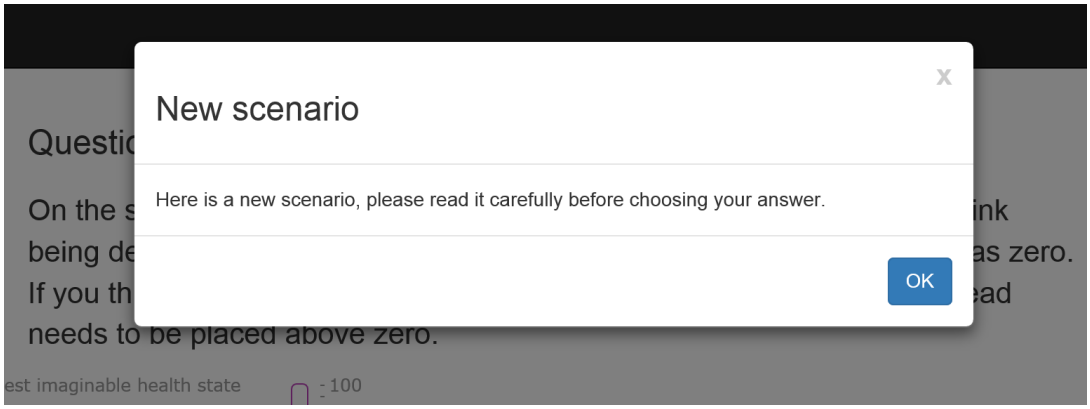
Survey company's details – to be added once confirmed.

If you agree to take part please click the continue button below.

Continue


ScHARR Survey
Process Utility survey
Participant Consent form
• Title of Project: Process Utility Survey
• Name of Researchers: Victoria Brennan, University of Sheffield
• Respondent ID for this project: AUTOMATIC
I confirm that I have read and understand the project information provided and if I require further details I can contact the researchers by email.
I understand that my participation is voluntary and that I am free to withdraw at any time without having to give a reason and without any sort of penalty.
I understand that any data I provide will be treated securely and kept confidential. My responses will be anonymised before analysis. Only the project team will have access to my responses.
I agree to take part in the Process Utility Survey.
I am a resident of the United Kingdom and I am over 18 years of age.
If you would like any further information about the research please contact the research team on ProcessSurvey@sheffield.ac.uk.
To continue to the survey please click Continue below, or click the Decline consent button if you do not wish to take part in the survey.
Decline consent
Continue

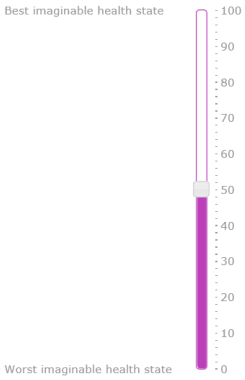

ScHARR Survey
Process Utility survey
Introduction
We will ask you to think about a set of 23 different states of health and tell us how good or bad you think they are. All of your responses will be treated as confidential, and all analysis will be carried out anonymously. We are interested in people's views, and there are no right or wrong answers. Please tell us what you think.
Continue


New scenario
Here is a new scenario, please read it carefully before choosing your answer.
OK
needs to be placed above zero.
100

ScHARR Survey

Question 1

On the scale shown below, where would you place being dead? If you think being dead is the worst imaginable health state then you would rate this as zero. If you think the worst imaginable health state is worse than dead, then dead needs to be placed above zero.



Save answer

ScHARR Survey

Question 2

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health does not limit you in moderate activities

You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems.

Your health limits your social activities none of the time

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You feel downhearted and low none of the time

You have a lot of energy most of the time

You have just had a consultation with the doctor. You evaluated your communication experience as very good

Save answer

ScHARR Survey

Question 3

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health does not limit you in moderate activities

You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems.

Your health limits your social activities none of the time

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You feel downhearted and low none of the time

You have a lot of energy most of the time

You have just had a consultation with the doctor. You evaluated your communication experience as good

Save answer

ScHARR Survey

Question 4

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health does not limit you in moderate activities

You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems.

Your health limits your social activities none of the time

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You feel downhearted and low none of the time

You have a lot of energy some of the time

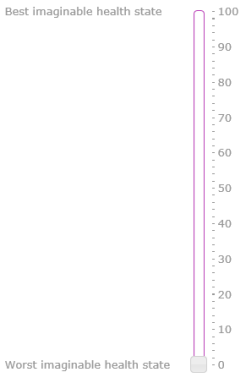
You have just had a consultation with the doctor. You evaluated your communication experience as very good

Save answer

ScHARR Survey

Question 5

On the scale between 0 and 100 shown, how would you SCORE this health state?



Your health does not limit you in moderate activities

You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems.

Your health limits your social activities none of the time

You feel downhearted and low none of the time

You have a lot of energy most of the time

You have pain that interferes with your normal work [both outside the home and housework] a little bit

You have just had a consultation with the doctor. You evaluated your communication experience as very good

Save answer

ScHARR Survey

Question 6

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health does not limit you in moderate activities

You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems.

Your health limits your social activities none of the time

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You have a lot of energy most of the time

You feel downhearted and low a little of the time

You have just had a consultation with the doctor. You evaluated your communication experience as average

Save answer

ScHARR Survey

Question 7

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health limits your social activities none of the time

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You feel downhearted and low none of the time

You have a lot of energy most of the time

Your health limits you a little in moderate activities

You are limited in the kind of work or other activities as a result of your physical health

You have just had a consultation with the doctor. You evaluated your communication experience as average

Save answer

ScHARR Survey

Question 8

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health does not limit you in moderate activities

Your health limits your social activities none of the time

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You have a lot of energy most of the time

You feel downhearted and low a little of the time

You are limited in the kind of work or other activities as a result of your physical health

You have just had a consultation with the doctor. You evaluated your communication experience as average

Save answer

ScHARR Survey

Question 9

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health does not limit you in moderate activities

Your health limits your social activities none of the time

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You have a lot of energy most of the time

You feel downhearted and low a little of the time

You are limited in the kind of work or other activities as a result of your physical health

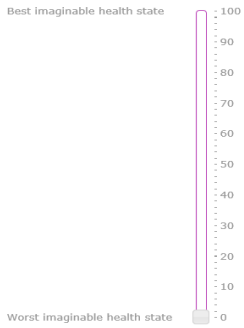
You have just had a consultation with the doctor. You evaluated your communication experience as poor

Save answer

ScHARR Survey

Question 10

On the scale between 0 and 100 shown, how would you SCORE this health state?



Your health limits your social activities none of the time

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You have a lot of energy most of the time

You feel downhearted and low a little of the time

Your health limits you a little in moderate activities

You are limited in the kind of work or other activities as a result of your physical health

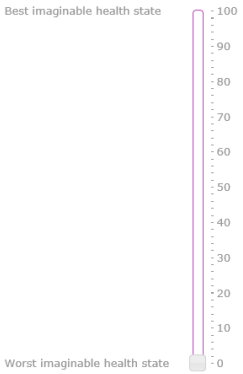
You have just had a consultation with the doctor. You evaluated your communication experience as good

Save answer

ScHARR Survey

Question 11

On the scale between 0 and 100 shown, how would you SCORE this health state?



Your health limits your social activities none of the time

You feel downhearted and low none of the time

You have a lot of energy most of the time

Your health limits you a little in moderate activities

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

You have pain that interferes with your normal work (both outside the home and housework) a little bit.

You have just had a consultation with the doctor. You evaluated your communication experience as very poor

Save answer

ScHARR Survey

Question 12

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health limits your social activities none of the time

You feel downhearted and low a little of the time

Your health limits you a little in moderate activities

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

You have pain that interferes with your normal work (both outside the home and housework) a little bit.

You have a lot of energy none of the time

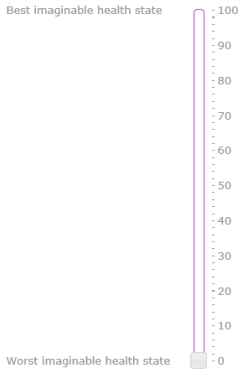
You have just had a consultation with the doctor. You evaluated your communication experience as very good

Save answer

ScHARR Survey

Question 13

On the scale between 0 and 100 shown, how would you SCORE this health state?



Your health does not limit you in moderate activities

Your health limits your social activities none of the time

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You accomplish less than you would like as a result of emotional problems.

You feel tense or downhearted and low some of the time

You have a lot of energy a most of the time

You have just had a consultation with the doctor. You evaluated your communication experience as poor

Save answer

ScHARR Survey

Question 14

On the scale between 0 and 100 shown, how would you SCORE this health state?

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You feel downhearted and low none of the time

You have a lot of energy most of the time

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

Your health limits you a lot in moderate activities

Your health limits your social activities some of the time

You have just had a consultation with the doctor. You evaluated your communication experience as average

Save answer

ScHARR Survey

Question 15

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health does not limit you in moderate activities

You have pain that does not interfere with your normal work (both outside the home and housework) at all

You have a lot of energy some of the time

You accomplish less than you would like as a result of emotional problems.

You feel tense or downhearted and low some of the time

Your health limits your social activities a little of the time

You have just had a consultation with the doctor. You evaluated your communication experience as poor

Save answer

ScHARR Survey

Question 16

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health does not limit you in moderate activities

You have a lot of energy some of the time

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

You have pain that interferes with your normal work (both outside the home and housework) a little bit.

Your health limits your social activities some of the time

You feel downhearted and low some of the time

You have just had a consultation with the doctor. You evaluated your communication experience as very poor

Save answer

ScHARR Survey

Question 17

On the scale between 0 and 100 shown, how would you SCORE this health state?

You have a lot of energy some of the time

You feel downhearted and low a little of the time

Your health limits you a little in moderate activities

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

You have pain that interferes with your normal work (both outside the home and housework) a little bit.

Your health limits your social activities some of the time

You have just had a consultation with the doctor. You evaluated your communication experience as good

Save answer

ScHARR Survey

Question 18

On the scale between 0 and 100 shown, how would you SCORE this health state?

Your health limits you a little in moderate activities

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

You have a lot of energy none of the time

You feel tense or downhearted and low some of the time

Your health limits your social activities some of the time

You have pain that interferes with your normal work (both outside the home and housework) moderately.

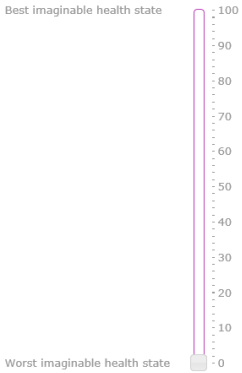
You have just had a consultation with the doctor. You evaluated your communication experience as good

Save answer

ScHARR Survey

Question 19

On the scale between 0 and 100 shown, how would you SCORE this health state?



Your health limits you a little in moderate activities

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

Your health limits your social activities some of the time

You feel downhearted and low some of the time

You have pain that interferes with your normal work (both outside the home and housework) quite a bit.

You have a lot of energy a little of the time

You have just had a consultation with the doctor. You evaluated your communication experience as very poor

Save answer

ScHARR Survey

Question 20

On the scale between 0 and 100 shown, how would you SCORE this health state?

You have a lot of energy some of the time

You feel downhearted and low a little of the time

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

Your health limits you a lot in moderate activities

Your health limits your social activities some of the time

You have pain that interferes with your normal work (both outside the home and housework) quite a bit.

You have just had a consultation with the doctor. You evaluated your communication experience as good

Save answer

ScHARR Survey

Question 21

On the scale between 0 and 100 shown, how would you SCORE this health state?

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

You have a lot of energy none of the time

Your health limits you a lot in moderate activities

Your health limits your social activities some of the time

You feel downhearted and low some of the time

You have pain that interferes with your normal work (both outside the home and housework) quite a bit.

You have just had a consultation with the doctor. You evaluated your communication experience as very poor

Save answer

ScHARR Survey

Question 22

On the scale between 0 and 100 shown, how would you SCORE this health state?

You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

You have a lot of energy none of the time

Your health limits you a lot in moderate activities

You feel downhearted and low some of the time

Your health limits your social activities all of the time

You have pain that interferes with your normal work (both outside the home and housework) extremely.

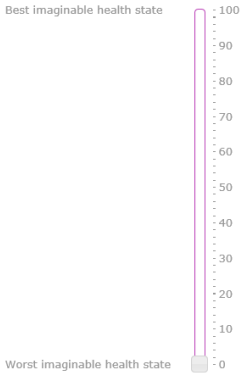
You have just had a consultation with the doctor. You evaluated your communication experience as poor

Save answer

ScHARR Survey

Question 23

On the scale between 0 and 100 shown, how would you SCORE this health state?



You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems.

You have a lot of energy none of the time

Your health limits you a lot in moderate activities

Your health limits your social activities all of the time

You have pain that interferes with your normal work (both outside the home and housework) extremely.

You feel downhearted and low all of the time

You have just had a consultation with the doctor. You evaluated your communication experience as very poor

Save answer

ScHARR Survey

Process Utility survey

PEQ Health States

We will now ask you to think about 25 scenarios describing your feelings immediately after a consultation with the doctor, your responses are highlighted in green. Please tell us if you would consider the communication experience to be very good, good, fair, poor or very poor.

Continue →

ScHARR Survey

Process Utility survey

Question 1

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 2

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 3

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 4

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 5

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 6

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 7

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 8

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 9

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 10

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 11

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 12

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

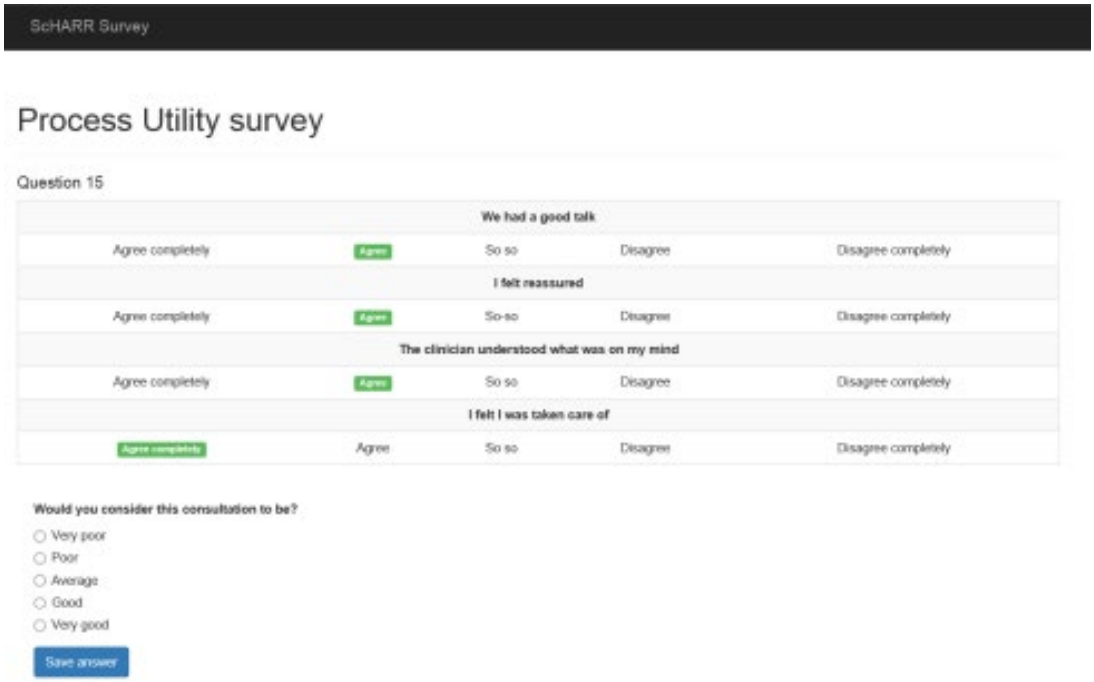
Question 13

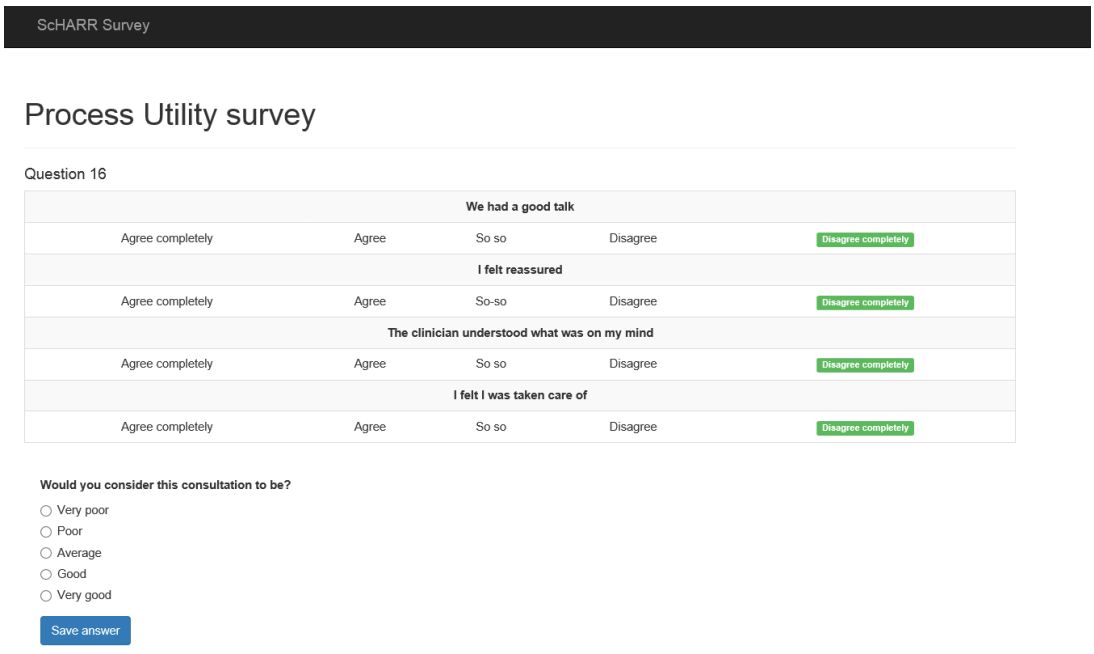
We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer


ScHARR Survey
Process Utility survey
Question 15
We had a good talk
Agree completely
Agree
So so
Disagree
Disagree completely
I felt reassured
Agree completely
Agree
So-so
Disagree
Disagree completely
The clinician understood what was on my mind
Agree completely
Agree
So so
Disagree
Disagree completely
I felt I was taken care of
Agree completely
Agree
So so
Disagree
Disagree completely
Would you consider this consultation to be?
Very poor
Poor
Average
Good
Very good
Save answer


ScHARR Survey
Process Utility survey
Question 16
We had a good talk
Agree completely
Agree
So so
Disagree
Disagree completely
I felt reassured
Agree completely
Agree
So-so
Disagree
Disagree completely
The clinician understood what was on my mind
Agree completely
Agree
So so
Disagree
Disagree completely
I felt I was taken care of
Agree completely
Agree
So so
Disagree
Disagree completely
Would you consider this consultation to be?
Very poor
Poor
Average
Good
Very good
Save answer

ScHARR Survey

Process Utility survey

Question 17

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 18

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 19

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 20

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 21

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 22

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 23

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 24

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Question 25

We had a good talk				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt reassured				
Agree completely	Agree	So-so	Disagree	Disagree completely
The clinician understood what was on my mind				
Agree completely	Agree	So so	Disagree	Disagree completely
I felt I was taken care of				
Agree completely	Agree	So so	Disagree	Disagree completely

Would you consider this consultation to be?

- ○ Very poor
- ○ Poor
- ○ Average
- ○ Good
- ○ Very good

Save answer

ScHARR Survey

Process Utility survey

Demographic questions

We will now ask you some questions about yourself

Continue →

ScHARR Survey

Process Utility survey

1. What is your age?

- ○ 18 - 24
- ○ 25 - 29
- ○ 30 - 34
- ○ 35 - 39
- ○ 40 - 44
- ○ 45 - 49
- ○ 50 - 54
- ○ 55 - 59
- ○ 60 - 64
- ○ 65 - 69
- ○ 70 - 74
- ○ 75+

Confirm

ScHARR Survey

Process Utility survey

2. What is your gender:

- ○ Male
- ○ Female

Confirm

ScHARR Survey

Process Utility survey

3. What is your ethnic origin:

- ○ White English/Welsh/Scottish/Northern Irish/British
- ○ White Irish
- ○ White Gypsy or Irish traveler
- ○ White Other
- ○ White and Black Carribean
- ○ White and Black African
- ○ White and Asian
- ○ Any other Mixed/multiple ethnic background
- ○ Indian
- ○ Pakistani
- ○ Bangladeshi
- ○ Chinese
- ○ Black/African/Caribbean/Black British African
- ○ Black/African/Caribbean/Black British Caribbean
- ○ Arab
- ○ Other

Confirm

ScHARR Survey

Process Utility survey

4. What is your highest education:

- ○ 1-4 O levels/CSEs/GCSEs (any grades), Enrty Level, Foundation Diploma
- ○ NVQ Level 1, Foundation GNVQ, Basic Skills
- ○ 5+ O levels (passes)/CSEs (grade 1)/GCSEs (Grades A*-C), School Certificate, 1 A level/2-3 AS levels/VCEs, Higher/Diploma
- ○ NVQ Level 2, Intermediate GNVQ, City and Guilds Craft, BTEC First/General Diploma, RSA Diploma
- ○ Apprenticeship
- ○ 2+ A levels / VCEs, 4+ AS levels, Higher School Certificate, Progression / Advanced Diploma
- ○ NVQ Level 3, Advanced GNVQ, City and Guilds Advanced Craft, ONC, OND, BTEC National, RSA Advanced Diplomas
- ○ Degree (for example BA, BSc), Higher degree (for example MA, PhD, PGCE)
- ○ NVQ Level 4 - 5, HNC, HND, RSA Higher Diploma, BTEC Higher Level
- ○ Professional qualifications (for example teaching, nursing, accountancy
- ○ Other vocational / work-related qualifications
- ○ Foreign qualifications
- ○ No qualifications

Confirm

ScHARR Survey

Process Utility survey

5. What is your current employment status?

- ○ Student
- ○ Employed
- ○ Home-maker
- ○ Self-employed or freelance
- ○ Unemployed
- ○ Long-term sick or disabled
- ○ Retired
- ○ Other

Confirm

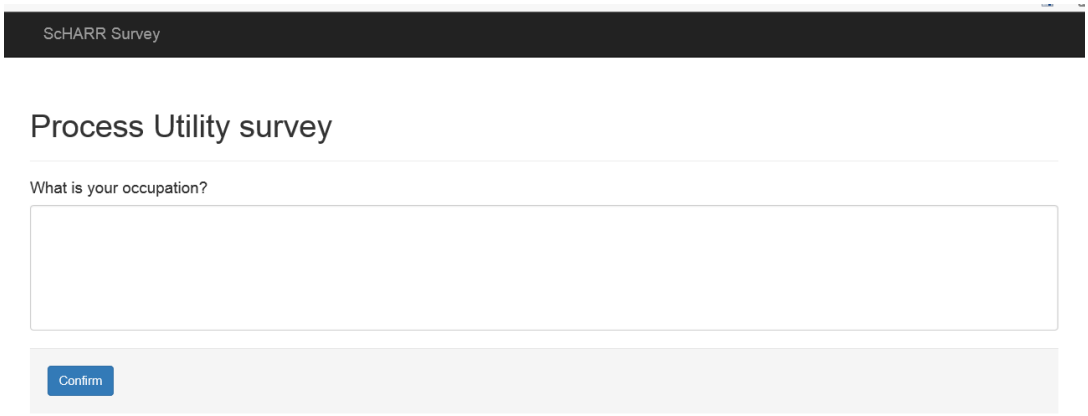
ScHARR Survey

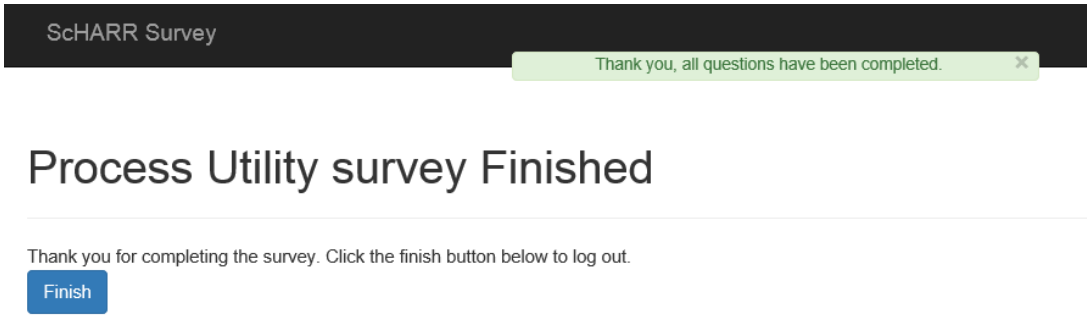
Process Utility survey

6. What is your marital status?

- ○ Never married and never registered in a same-sex civil partnership
- ○ Married/In a registered same-sex civil partnership
- ○ Separated, but still legally married/separated but still legally in a same-sex civil partnership
- ○ Divorced/formerly in a same-sex civil partnership which is now legally dissolved
- ○ Widowed/surviving partner from a same-sex civil partnership

Confirm


ScHARR Survey
Process Utility survey
What is your occupation?
Confirm


ScHARR Survey
Thank you, all questions have been completed.
Process Utility survey Finished
Thank you for completing the survey. Click the finish button below to log out.
Finish

Appendix 25: Multiple Linear Regression Results (basecase)

This appendix contains the output for the regressions. Multiple linear regression was used to assess the association of the process domain and mean utility across the study sample. Regression was run on mean utility.

Three separate specifications were assessed: the first estimated mean utility as a function of a constant plus process level, the second also included a measure of overall health state severity, and the third included health state severity, plus an interaction term between health state severity and process level. The model for the latter, and from which the other two models are simplifications, is defined as:

Mean utility = α + process level + severity + process level * severity + ε

Where the process level is defined as a categorical variable with 3 levels (poor, fair, good) and severity is defined by characteristics of the underlying SF-6D health state. In order to simplify the analysis, in the presence of only 22 data points, severity was represented with dummy variables, where 1 = healthy patients, and 0 = sick patients. These are binary data categories determined as values falling above and below the median SF-6D utility scores obtained from the survey.

Table 41. Mean process level decrements (3 levels).

Process utility decrement level	N	Minimum	Maximum	Mean	Std. Deviation
Good	9	-7.48	-.94	-4.5331	2.44675
Fair	4	-7.63	-2.48	-5.7829	2.39775
Poor	9	-11.42	-1.97	-7.0989	3.33829

Model 1: Mean utility as a function of a constant plus process level

Table 42. Model 1 Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.401[a]	.161	.119	2.77757

a. Predictors: (Constant), Process_Domain

Table 43. Model 1 ANOVA[a]

	Sum of Squares	df	Mean Square	F	Sig.
Regression	29.625	1	29.625	3.840	.064[b]
Residual	154.297	20	7.715		
Total	183.922	21			

a. Dependent Variable: PU_Decrement

b. Predictors: (Constant), Process_Domain

Table 44. Model 1 Co-efficients[a]

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	-3.244	1.437		-2.258	.035
Process domain	-1.283	.655	-.401	-1.960	.064

a. Dependent Variable: PU_Decrement

Model 2: Mean utility as a function of a constant plus process level plus severity

Table 45. Model 2 Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
2	.773[a]	.598	.555	1.97362

a. Predictors: (Constant), Severity, Process_Domain

Table 46. Model 2 ANOVA[a]

	Sum of Squares	df	Mean Square	F	Sig.
Regression	109.914	2	54.957	14.109	.000[b]
Residual	74.009	19	3.895		
Total	183.922	21			

a. Dependent Variable: PU_Decrement

b. Predictors: (Constant), Severity, Process_Domain

Table 47. Model 2 Co-efficients[a]

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	-.427	1.195		-.358	.725
Process domain	-1.716	.475	-.537	-3.614	.002
Severity	-3.900	.859	-.674	-4.540	.000

a. Dependent Variable: PU_Decrement

Model 3: Mean utility as a function of a constant plus process level plus severity, plus interaction (process level*severity).

Table 48. Model 3 Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
3	.773[a]	.598	.531	2.02767

a. Predictors: (Constant), Interaction, Severity, Process_Domain

Table 49. Model 3 ANOVA[a]

	Sum of Squares	df	Mean Square	F	Sig.
Regression	109.917	3	36.639	8.911	.001b
Residual	74.006	18	4.111		
Total	183.922	21			

a. Dependent Variable: PU_Decrement

b. Predictors: (Constant), Interaction, Severity, Process_Domain

Table 50. Model 3 Co-efficients[a]

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	-.402	1.551		-.259	.798
Process domain	-1.728	.653	-.540	-2.645	.016
Severity	-3.952	2.135	-.683	-1.851	.081
Interaction	.026	.982	.010	.026	.979

a. Dependent Variable: PU_Decrement

Appendix 26: Multiple Linear Regression Results (sensitivity analysis)

Model 1: Mean utility as a function of a constant plus process level

Table 51. Model 1 Summary (Sensitivity analysis)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.518[a]	.268	.231	9.79502

a. Predictors: (Constant), Process_Domain

Table 52. Model 1 ANOVA[a] (Sensitivity analysis)

	Sum of Squares	df	Mean Square	F	Sig.
Regression	702.625	1	702.625	7.323	.014
Residual	1918.846	20	95.942		
Total	2621.472	21			

a. Dependent Variable: PU_Decrement

b. Predictors: (Constant), Process_Domain

Table 53. Model 1 Co-efficients[a] (Sensitivity analysis)

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	-.864	5.068		-.170	.866
Process domain	-6.248	2.309	-.518	-2.706	.014

a. Dependent Variable: PU_Decrement

Model 2: Mean utility as a function of a constant plus process level plus severity

Table 54. Model 2 Summary (Sensitivity analysis)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
2	.519[a]	.269	.192	10.04006

a. Predictors: (Constant), Severity, Process_Domain

Table 55. Model 2 ANOVA[a] (Sensitivity analysis)

	Sum of Squares	df	Mean Square	F	Sig.
Regression	706.219	2	353.110	3.503	.051[b]
Residual	1915.252	19	100.803		
Total	2621.472	21			

a. Dependent Variable: PU_Decrement

b. Predictors: (Constant), Severity, Process_Domain

Table 56. Model 2 Co-efficients[a] (Sensitivity analysis)

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	-1.795	7.165		-.251	.805
Process domain	-6.444	2.585	-.534	-2.493	.022
Severity	.883	4.676	.040	.189	.852

a. Dependent Variable: PU_Decrement

Model 3: Mean utility as a function of a constant plus process level plus severity, plus interaction (process level*severity).

Table 57. Model 3 Summary (Sensitivity analysis)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
3	.521[a]	.272	.150	10.30019

a. Predictors: (Constant), Interaction, Severity, Process_Domain

Table 58. Model 3 ANOVA[a] (Sensitivity analysis)

	Sum of Squares	df	Mean Square	F	Sig.
Regression	711.781	3	237.260	2.236	.119[b]
Residual	1909.690	18	106.094		
Total	2621.472	21			

a. Dependent Variable: PU_Decrement

b. Predictors: (Constant), Interaction, Severity, Process_Domain

Table 59. Model 3 Co-efficients[a] (Sensitivity analysis)

	Unstandardized Coefficients		Standard-ized Coeffi-cients		
	B	Std. Er-ror	Beta	t	Sig.
(Constant)	1.731	17.064		.101	.920
Process domain	-8.363	8.789	-.693	-.951	.354
Severity	1.225	5.350	.231	.229	.821
Interaction	-1.508	11.491	-.069	-.131	.897

a. Dependent Variable: PU_Decrement

Appendix 27: Multinomial Regression Results

This appendix contains the output from the multinomial regression which was used to determine which combinations of the PEQ communication responses from the trial correspond with which rating score (Good, fair or poor). The categorical predictors (factors) were the PEQ responses. There were 4 predictors for each rating scale (4 questions within the PEQ communication domain). The dependent variable was the rating score. There were no covariates, and a main effects model was run so as to include both main effects and factor-by-factor interactions.

Table 60. Model Fitting information

	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC	BIC	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	5775.480	5802.772	5767.480			
Final	955.872	1419.830	819.872	4947.608	64	.000

The Log-likelihood shown in Table 60 is a measure of how much unexplained variability there was in the data; therefore the difference or change in log-likelihood indicates how much new variance has been explained by the model.

Chi-square tests the decrease in unexplained variance from the baseline model (5767.480) to the final model (819.872), which is a difference of 4947.608. This chi-square statistic for the main effects model is significant (<0.05), indicating that factor by factor interactions (interactions between the PEQ responses) have a significant effect on predicting the rating of the health state.

It has been ascertained that the model is significantly better than no model, but not whether or not it is a good fit of the data. The Pearson and Deviance tests, reported in Table 61 ascertain whether predicted values from the model differ significantly

from the observed values. Lack of significance can be interpreted as indicating good fit (p-value >0.05). Both tests have a significance <0.01, therefore indicating that the predicted values are significantly different from the observed values.

Table 61. Goodness-of-fit.

	Chi-Square	df	Sig.
Pearson	5505.047	32	.000
Deviance	376.489	32	.000

Calculation of the R^2 shown in Table 62, is an additional measure of the goodness of fit, and is based on the log-likelihood of the new model and the log-likelihood of the original model, and the sample size. The outcome can range from 0 (indicating that the predictors are not good at predicting the outcome variable), to 1 (indicating that the model predicts the outcome perfectly). These results therefore indicate an average fit to the model.

Table 62. Pseudo R-square

Measure	Value
Cox and Snell	.518
Nagelkerke	.547
McFadden	.250

There are two outputs related to the statistical significance of individual predictor variables: the Likelihood Ratio Tests and Parameter Estimates. The results of the likelihood ratio tests can be used to ascertain the significance of predictors (PEQ responses) to the model. The results in Table 63 show that all are significant, and therefore can be used to predict the overall rating of the health states. These are overall statistics.

Table 63. Likelihood Ratio Tests

	Model Fitting Criteria			Likelihood Ratio Tests		
	AIC of reduced model	BIC of reduced model	-2 Log Likelihood of reduced model	Chi-Square	df	Sig.
Intercept	955.872	1419.830	819.872[a]	.000	0	.
PEQ_comm_1	1532.218	1887.009	1428.218	608.346	16	.000
PEQ_comm_2	1807.466	2162.257	1703.466	883.594	16	.000
PEQ_comm_3	1872.850	2227.642	1768.850	948.978	16	.000
PEQ_comm_4	2259.492	2614.283	2155.492	1335.620	16	.000

Note: The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

This analysis resulted in the table of parameter estimates in Table 64. The co-efficient's calculated through running the MNR analysis on the survey data were then used on the ePAQ trial data to estimate the probability of membership into each rating score group, based on the combinations participants PEQ responses.

Table 64. Multinomial regression Parameter estimates

Rating for health state[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
Very poor	Intercept	-8.667	.487	316.187	1	.000			
	[PEQ_comm_1=1]	3.937	.349	126.882	1	.000	51.251	25.835	101.668
	[PEQ_comm_1=2]	4.017	.352	130.436	1	.000	55.546	27.877	110.676
	[PEQ_comm_1=3]	1.438	.333	18.583	1	.000	4.211	2.190	8.095
	[PEQ_comm_1=4]	.500	.315	2.523	1	.112	1.649	.890	3.055
	[PEQ_comm_1=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_2=1]	4.215	.370	129.468	1	.000	67.714	32.760	139.964
	[PEQ_comm_2=2]	3.794	.357	112.760	1	.000	44.421	22.054	89.472
	[PEQ_comm_2=3]	3.145	.326	92.981	1	.000	23.213	12.250	43.988
	[PEQ_comm_2=4]	-.482	.297	2.635	1	.105	.617	.345	1.105

Rating for health state[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
	[PEQ_comm_2=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_3=1]	4.880	.348	197.058	1	.000	131.674	66.615	260.272
	[PEQ_comm_3=2]	3.894	.357	118.954	1	.000	49.108	24.392	98.870
	[PEQ_comm_3=3]	3.115	.380	67.151	1	.000	22.545	10.701	47.497
	[PEQ_comm_3=4]	.281	.310	.821	1	.365	1.324	.721	2.431
	[PEQ_comm_3=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_4=1]	5.947	.391	231.221	1	.000	382.756	177.827	823.848
	[PEQ_comm_4=2]	3.851	.347	122.934	1	.000	47.037	23.812	92.913
	[PEQ_comm_4=3]	1.783	.331	29.025	1	.000	5.946	3.109	11.374
	[PEQ_comm_4=4]	1.901	.357	28.315	1	.000	6.692	3.323	13.478
	[PEQ_comm_4=5]	0b	.	.	0	.	.	.	.
Poor	Intercept	-3.923	.300	171.064	1	.000			

Rating for health state[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
	[PEQ_comm_1=1]	2.102	.286	54.027	1	.000	8.186	4.673	14.339
	[PEQ_comm_1=2]	2.519	.303	69.275	1	.000	12.413	6.859	22.462
	[PEQ_comm_1=3]	1.159	.247	22.021	1	.000	3.187	1.964	5.171
	[PEQ_comm_1=4]	.124	.231	.289	1	.591	1.132	.720	1.782
	[PEQ_comm_1=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_2=1]	2.994	.304	96.961	1	.000	19.972	11.005	36.247
	[PEQ_comm_2=2]	2.935	.290	102.428	1	.000	18.824	10.662	33.233
	[PEQ_comm_2=3]	2.212	.270	66.914	1	.000	9.129	5.374	15.508
	[PEQ_comm_2=4]	-.130	.228	.324	1	.569	.878	.562	1.373
	[PEQ_comm_2=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_3=1]	3.027	.282	115.059	1	.000	20.634	11.868	35.874
	[PEQ_comm_3=2]	2.808	.278	102.125	1	.000	16.578	9.616	28.580

Rating for health state[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
	[PEQ_comm_3=3]	2.054	.305	45.327	1	.000	7.802	4.290	14.189
	[PEQ_comm_3=4]	.217	.231	.885	1	.347	1.242	.791	1.952
	[PEQ_comm_3=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_4=1]	4.024	.323	154.962	1	.000	55.951	29.691	105.438
	[PEQ_comm_4=2]	2.777	.259	115.379	1	.000	16.066	9.680	26.665
	[PEQ_comm_4=3]	1.435	.236	36.924	1	.000	4.201	2.644	6.674
	[PEQ_comm_4=4]	1.196	.256	21.833	1	.000	3.306	2.002	5.458
	[PEQ_comm_4=5]	0b	.	.	0	.	.	.	.
Average	Intercept	-.905	.246	13.556	1	.000			
	[PEQ_comm_1=1]	1.199	.274	19.164	1	.000	3.315	1.938	5.669
	[PEQ_comm_1=2]	1.138	.296	14.786	1	.000	3.119	1.747	5.570
	[PEQ_comm_1=3]	.682	.237	8.241	1	.004	1.977	1.241	3.149

Rating for health state[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
	[PEQ_comm_1=4]	-.305	.215	2.017	1	.156	.737	.484	1.123
	[PEQ_comm_1=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_2=1]	1.972	.285	47.953	1	.000	7.187	4.113	12.560
	[PEQ_comm_2=2]	1.502	.281	28.481	1	.000	4.490	2.586	7.794
	[PEQ_comm_2=3]	1.603	.256	39.299	1	.000	4.967	3.009	8.198
	[PEQ_comm_2=4]	-.658	.209	9.933	1	.002	.518	.344	.780
	[PEQ_comm_2=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_3=1]	2.032	.263	59.687	1	.000	7.631	4.557	12.779
	[PEQ_comm_3=2]	1.600	.267	35.913	1	.000	4.952	2.935	8.357
	[PEQ_comm_3=3]	2.126	.291	53.279	1	.000	8.382	4.736	14.836
	[PEQ_comm_3=4]	.413	.206	4.010	1	.045	1.511	1.009	2.264
	[PEQ_comm_3=5]	0b	.	.	0	.	.	.	.

Rating for health state[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
	[PEQ_comm_4=1]	2.063	.308	44.727	1	.000	7.870	4.299	14.406
	[PEQ_comm_4=2]	.948	.247	14.668	1	.000	2.580	1.588	4.191
	[PEQ_comm_4=3]	.357	.217	2.704	1	.100	1.430	.934	2.189
	[PEQ_comm_4=4]	1.087	.234	21.664	1	.000	2.965	1.876	4.687
	[PEQ_comm_4=5]	0b	.	.	0	.	.	.	.
Good	Intercept	.713	.250	8.154	1	.004			
	[PEQ_comm_1=1]	.902	.290	9.681	1	.002	2.464	1.396	4.349
	[PEQ_comm_1=2]	1.061	.313	11.510	1	.001	2.888	1.565	5.330
	[PEQ_comm_1=3]	.298	.257	1.349	1	.246	1.347	.815	2.227
	[PEQ_comm_1=4]	-.035	.234	.022	1	.882	.966	.611	1.527
	[PEQ_comm_1=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_2=1]	.361	.300	1.448	1	.229	1.435	.797	2.583

Rating for health state[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
	[PEQ_comm_2=2]	-.034	.311	.012	1	.914	.967	.526	1.778
	[PEQ_comm_2=3]	.396	.271	2.141	1	.143	1.486	.874	2.525
	[PEQ_comm_2=4]	-.748	.221	11.448	1	.001	.473	.307	.730
	[PEQ_comm_2=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_3=1]	-.074	.284	.069	1	.793	.928	.532	1.619
	[PEQ_comm_3=2]	.007	.287	.001	1	.981	1.007	.574	1.767
	[PEQ_comm_3=3]	.647	.307	4.448	1	.035	1.910	1.047	3.485
	[PEQ_comm_3=4]	-.464	.212	4.811	1	.028	.629	.415	.952
	[PEQ_comm_3=5]	0b	.	.	0	.	.	.	.
	[PEQ_comm_4=1]	.412	.332	1.544	1	.214	1.510	.788	2.892
	[PEQ_comm_4=2]	-.398	.280	2.019	1	.155	.672	.388	1.163
	[PEQ_comm_4=3]	-.161	.232	.482	1	.488	.851	.540	1.342

Rating for health state[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
	[PEQ_comm_4=4]	1.238	.237	27.400	1	.000	3.450	2.170	5.485
	[PEQ_comm_4=5]	0b	.	.	0	.	.	.	.

a. The reference category is: Very good.

b. This parameter is set to zero because it is redundant.

Table 65. Observed and predicted frequencies.

Response to PEQ 4	Response to PEQ 3	Response to PEQ 2	Response to PEQ 1	Rating for health state	Frequency			Percentage	
					Observed	Predicted	Pearson Residual	Observed	Predicted
Disagree completely	Disagree completely	Disagree completely	Disagree completely	Very poor	242	236.962	.924	89.3%	87.4%
				Poor	8	29.394	-4.179	3.0%	10.8%
				Average	8	4.557	1.627	3.0%	1.7%
				Good	7	.080	24.541	2.6%	0.0%
				Very good	6	.008	67.545	2.2%	0.0%
Disagree completely	Disagree	Agree completely	So-so	Very poor	26	33.913	-1.453	9.6%	12.6%
				Poor	152	145.598	.782	56.3%	53.9%
				Average	79	77.606	.188	29.3%	28.7%
				Good	10	10.394	-.125	3.7%	3.8%
				Very good	3	2.489	.325	1.1%	0.9%
Disagree completely	So-so	Agree	Agree completely	Very poor	14	9.591	1.449	5.1%	3.5%
				Poor	80	79.338	.088	29.3%	29.1%
				Average	152	144.513	.908	55.7%	52.9%

Response to PEQ 4	Response to PEQ 3	Response to PEQ 2	Response to PEQ 1	Rating for health state	Frequency			Percentage	
					Observed	Predicted	Pearson Residual	Observed	Predicted
				Good	24	29.101	-1.000	8.8%	10.7%
				Very good	3	10.457	-2.352	1.1%	3.8%
Disagree completely	Agree	So-so	Disagree	Very poor	85	86.238	-.162	31.5%	31.9%
				Poor	132	119.474	1.535	48.9%	44.2%
				Average	39	57.153	-2.704	14.4%	21.2%
				Good	10	6.368	1.456	3.7%	2.4%
				Very good	4	.767	3.698	1.5%	0.3%
Disagree completely	Agree completely	Disagree	Agree	Very poor	72	66.172	.825	26.8%	24.6%
				Poor	13	18.057	-1.232	4.8%	6.7%
				Average	4	6.279	-.920	1.5%	2.3%
				Good	35	27.554	1.497	12.9%	10.2%
				Very good	124	116.455	.926	45.8%	43.0%
Disagree	Disagree completely	Agree completely	Agree	Very poor	35	27.554	1.497	12.9%	10.2%
				Poor	124	116.455	.926	45.8%	43.0%
				Average	88	92.055	-.520	32.5%	34.0%
				Good	16	19.259	-.770	5.9%	7.1%
				Very good	8	15.677	-1.998	3.0%	5.8%

Response to PEQ 4	Response to PEQ 3	Response to PEQ 2	Response to PEQ 1	Rating for health state	Frequency			Percentage	
					Observed	Predicted	Pearson Residual	Observed	Predicted
Disagree	Disagree	Agree	Disagree completely	Very poor	51	55.444	-.669	18.7%	20.3%
				Poor	178	166.962	1.371	65.2%	61.2%
				Average	30	39.101	-1.572	11.0%	14.3%
				Good	7	7.086	-.033	2.6%	2.6%
				Very good	7	4.406	1.246	2.6%	1.6%
Disagree	So-so	So-so	So-so	Very poor	33	26.935	1.231	12.0%	9.8%
				Poor	95	108.937	-1.718	34.5%	39.6%
				Average	126	129.715	-.449	45.8%	47.2%
				Good	11	7.905	1.117	4.0%	2.9%
				Very good	10	1.510	6.929	3.6%	0.5%
Disagree	Agree	Disagree	Agree completely	Very poor	5	7.840	-1.029	1.8%	2.8%
				Poor	107	122.373	-1.860	38.6%	44.2%
				Average	142	116.614	3.089	51.3%	42.1%
				Good	13	13.710	-.197	4.7%	4.9%
				Very good	10	16.462	-1.642	3.6%	5.9%
Disagree	Agree completely	Disagree completely	Disagree	Very poor	53	59.227	-.915	19.6%	21.9%
				Poor	164	153.273	1.315	60.5%	56.6%

Response to PEQ 4	Response to PEQ 3	Response to PEQ 2	Response to PEQ 1	Rating for health state	Frequency			Percentage	
					Observed	Predicted	Pearson Residual	Observed	Predicted
				Average	37	45.515	-1.384	13.7%	16.8%
				Good	12	11.040	.295	4.4%	4.1%
				Very good	5	1.945	2.199	1.8%	0.7%
So-so	Disagree completely	Agree	Disagree	Very poor	35	36.813	-.322	13.1%	13.7%
				Poor	141	148.915	-.973	52.6%	55.6%
				Average	75	56.772	2.725	28.0%	21.2%
				Good	13	17.535	-1.120	4.9%	6.5%
				Very good	4	7.964	-1.426	1.5%	3.0%
So-so	Disagree	So-so	Agree	Very poor	16	17.154	-.288	5.9%	6.4%
				Poor	116	127.048	-1.349	43.1%	47.2%
				Average	109	93.523	1.981	40.5%	34.8%
				Good	23	22.356	.142	8.6%	8.3%
				Very good	5	8.918	-1.334	1.9%	3.3%
So-so	So-so	Disagree	Disagree completely	Very poor	58	61.190	-.463	21.3%	22.5%
				Poor	134	116.390	2.158	49.3%	42.8%
				Average	64	84.058	-2.632	23.5%	30.9%
				Good	12	9.197	.940	4.4%	3.4%

Response to PEQ 4	Response to PEQ 3	Response to PEQ 2	Response to PEQ 1	Rating for health state	Frequency			Percentage	
					Observed	Predicted	Pearson Residual	Observed	Predicted
				Very good	4	1.165	2.633	1.5%	0.4%
So-so	Agree	Disagree completely	So-so	Very poor	9	4.760	1.960	3.2%	1.7%
				Poor	87	80.943	.800	31.4%	29.2%
				Average	158	153.003	.604	57.0%	55.2%
				Good	19	25.977	-1.438	6.9%	9.4%
				Very good	4	12.317	-2.424	1.4%	4.4%
So-so	Agree completely	Agree completely	Agree completely	Very poor	2	.083	6.675	0.7%	0.0%
				Poor	2	6.703	-1.839	0.7%	2.4%
				Average	28	46.644	-2.997	10.2%	17.0%
				Good	148	139.934	.975	54.0%	51.1%
				Very good	94	80.637	1.771	34.3%	29.4%
Agree	Disagree completely	So-so	Agree completely	Very poor	7	13.396	-1.792	2.6%	4.9%
				Poor	63	46.879	2.586	23.0%	17.1%
				Average	162	173.005	-1.378	59.1%	63.1%
				Good	36	36.915	-.162	13.1%	13.5%
				Very good	6	3.805	1.133	2.2%	1.4%
Agree	Disagree	Disagree	Disagree	Very poor	87	76.398	1.430	32.0%	28.1%

Response to PEQ 4	Response to PEQ 3	Response to PEQ 2	Response to PEQ 1	Rating for health state	Frequency			Percentage	
					Observed	Predicted	Pearson Residual	Observed	Predicted
				Poor	129	138.684	-1.175	47.4%	51.0%
				Average	41	45.546	-.738	15.1%	16.7%
				Good	12	10.825	.365	4.4%	4.0%
				Very good	3	.547	3.319	1.1%	0.2%
Agree	So-so	Disagree completely	Agree	Very poor	3	8.961	-2.025	1.1%	3.3%
				Poor	37	35.681	.237	13.7%	13.2%
				Average	168	164.701	.412	62.2%	61.0%
				Good	58	57.565	.065	21.5%	21.3%
				Very good	4	3.091	.520	1.5%	1.1%
Agree	Agree	Agree completely	Disagree completely	Very poor	1	1.127	-.120	0.4%	0.4%
				Poor	6	9.591	-1.181	2.2%	3.6%
				Average	95	86.681	1.085	35.3%	32.2%
				Good	154	157.181	-.394	57.2%	58.4%
				Very good	13	14.419	-.384	4.8%	5.4%
Agree	Agree completely	Agree	So-so	Very poor	2	.117	5.501	0.7%	0.0%
				Poor	3	7.165	-1.577	1.1%	2.7%
				Average	52	48.066	.626	19.3%	17.8%

Response to PEQ 4	Response to PEQ 3	Response to PEQ 2	Response to PEQ 1	Rating for health state	Frequency			Percentage	
					Observed	Predicted	Pearson Residual	Observed	Predicted
				Good	178	175.514	.317	65.9%	65.0%
				Very good	35	39.138	-.715	13.0%	14.5%
Agree completely	Disagree completely	Disagree	So-so	Very poor	15	19.275	-1.010	5.5%	7.1%
				Poor	117	111.357	.697	43.2%	41.1%
				Average	118	124.610	-.806	43.5%	46.0%
				Good	13	11.211	.546	4.8%	4.1%
				Very good	8	4.547	1.633	3.0%	1.7%
Agree completely	Disagree	Disagree completely	Agree completely	Very poor	9	6.090	1.193	3.3%	2.2%
				Poor	73	69.708	.458	26.9%	25.7%
				Average	150	153.224	-.395	55.4%	56.5%
				Good	30	31.339	-.254	11.1%	11.6%
				Very good	9	10.639	-.513	3.3%	3.9%
Agree completely	So-so	Agree completely	Disagree	Very poor	1	2.323	-.872	0.4%	0.9%
				Poor	15	20.654	-1.295	5.6%	7.7%
				Average	127	114.014	1.602	47.2%	42.4%
				Good	120	121.232	-.151	44.6%	45.1%
				Very good	6	10.777	-1.485	2.2%	4.0%

Response to PEQ 4	Response to PEQ 3	Response to PEQ 2	Response to PEQ 1	Rating for health state	Frequency			Percentage	
					Observed	Predicted	Pearson Residual	Observed	Predicted
Agree completely	Agree	Agree	Agree	Very poor	0	.034	-.185	0.0%	0.0%
				Poor	4	3.619	.201	1.5%	1.3%
				Average	14	34.548	-3.741	5.1%	12.7%
				Good	94	86.763	.941	34.4%	31.8%
				Very good	161	148.035	1.575	59.0%	54.2%
Agree completely	Agree completely	So-so	Disagree completely	Very poor	6	3.277	1.513	2.2%	1.2%
				Poor	20	23.662	-.788	7.4%	8.8%
				Average	124	106.603	2.169	46.1%	39.6%
				Good	113	119.455	-.792	42.0%	44.4%
				Very good	6	16.002	-2.578	2.2%	5.9%

Appendix 28: Process Utility Think Aloud Study ScHARR Ethics Application

This form has been approved by the University Research Ethics Committee (UREC)

Date:	13th February 2014
Name of applicant:	Victoria Brennan
Research project title:	A “think aloud” study: Investigating how people respond to SF-7D health states.

Complete this form if you are a **member of staff or a postgraduate research student** who plans to undertake a research project which requires ethics approval via the University Ethics Review Procedure. | X

or

Complete this form if you plan to submit a **'generic' research ethics application (i.e. an application** that will cover several sufficiently similar research projects). Information on the 'generic' route is at: www.sheffield.ac.uk/ris/other/gov-ethics/ethicspolicy/approval-procedure/review-procedure/generic-research-projects | |

If you are an undergraduate or a postgraduate-taught student, this is the wrong form.

This form should be accompanied, where appropriate, by all Information Sheets/Covering Letters/Written Scripts which you propose to use to inform the prospective participants about the proposed research, and/or by a Consent Form where you need to use one.

Further guidance on how to apply is at: http://www.shef.ac.uk/scharr/research/ethicsgovernance

Guidance on the possible routes for obtaining ethics approval (i.e. on the University Ethics Review Procedure, the NHS procedure and the Social Care Research Ethics Committee, and the Alternative procedure) is at: www.shef.ac.uk/ris/other/gov-ethics/ethicspolicy/approval-procedure/ethics-approval

Once you have completed this research ethics application form in full, and other documents where appropriate, check that your name, the title of your research project and the date is contained in the footer of each page and email, as a word document, to the Ethics Administrator k.woodhead@sheffield.ac.uk. Please note that the original signed and dated version of 'Part B' of the application form should be provided to the Ethics Administrator in hard copy.

I confirm that I have read the current version of the University of Sheffield 'Ethics Policy Governing Research Involving Human Participants, Personal Data and Human Tissue', as shown on the University's research ethics website at: www.shef.ac.uk/ris/other/gov-ethics/ethicspolicy Y

Part A

A1. Title of Research Project: A "think aloud" study: Investigating how people respond to SF-7D health states.

A2. Contact person (normally the Principal Investigator, in the case of staff-led research projects, or the student in the case of supervised-postgraduate researcher projects):

Title: Miss

Name: Victoria Brennan

Post: PhD student

Department: HEDS, ScHARR

Email: vkbrennan@hotmail.com

Telephone: 07725010182

A2.1. Is this a postgraduate researcher project? If yes, please provide the Supervisor's contact details:

Title: Dr

Post: Reader in Social Science

Email: g.l.jones@sheffield.ac.uk

Name: Georgina Jones

Department: ScHARR

Telephone: 01142 220749

A2.2. Other key investigators/co-applicants (within/outside University), where applicable. Please list all (add more if necessary):

Nil

A3. Proposed Project Duration:

Start date: March 2014 End date: December 2014

A4. Mark 'X' in one or more of the following boxes if your research:

	involves adults with mental incapacity or mental illness
	involves prisoners or others in custodial care (e.g. young offenders)
	involves children or young people aged under 18 years
	involves using samples of human biological material collected before for another purpose
	involves taking new samples of human biological material (e.g. blood, tissue) *
	involves testing a medicinal product *
	involves taking new samples of human biological material (e.g. blood, tissue) *
	involves additional radiation above that required for clinical care *

	involves investigating a medical device *

	is social care research

	is ESRC funded
	Is taking place in the health service but does not require NHS ethical approval**
	URMS number if required (please see below)

* If you have marked boxes marked ***** then you also need to obtain confirmation that appropriate University insurance is in place. The procedure for doing so is entirely by email. Please send an email addressed to insurance@shef.ac.uk and request a copy of the 'Clinical Trial Insurance Application Form'.

- If you have marked the box** your supervisor, needs to obtain an URMS number (details on the ScHARR web site http://www.shef.ac.uk/scharr/research/ethicsgovernance/ugpgt)

It is recommended that you familiarise yourself with the University's Ethics Policy Governing Research Involving Human Participants, Personal Data and Human Tissue before completing the following questions. Please note that if you provide sufficient information about the research (what you intend to do, how it will be carried out and how you intend to minimise any risks), this will help the ethics reviewers to make an informed judgement quickly without having to ask for further details.

A5. Briefly summarise:

iii. The project's aims and objectives:
(this must be in language comprehensible to a lay person)

To improve decision-making about funding different health care programmes health economists aim to measure and compare the benefits of different health care treatments. These benefits are often measured by asking the general public their views about the desirability of different health state descriptions.

Health economists often use 'choice tasks' to measure peoples' desire for different health state descriptions. These choice tasks are methods such as the 'Visual Analogue Scale', the 'Time Trade Off' and the 'Standard Gamble'. During choice tasks people are presented with a health outcome and asked a series of questions. For example, in the time trade off participants decide between living a certain number of years in 'less than Full Health' or a lesser number of years in 'Full Health'. By discovering how many years people are willing to 'sacrifice', health economists can infer the desirability of different health outcomes.

Some research does exist which explores the thought processes people have in completing these choice tasks. For example, Baker and Robinson (2004) presented respondents with health state descriptions related to hypertension and asked them to value these using a technique called standard gamble; and Spencer (2001) explored issues relating to the use of both standard gamble and time-trade off techniques in health state

valuations. The health state descriptions used in these exercises most commonly include only health outcomes. For example, a study performed by Baker and Robinson (2004) included the following health state description:

1. Your arm and leg are a little weak on one side
2. Your speech is a little slurred but people understand you
3. You may be unable to perform some of your usual activities
4. You can look after yourself as usual

However, a recent study performed at Sheffield University also included information related to the process of receiving health care:

1. Your health does not limit you in moderate activities
2. You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems.
3. Your health limits your social activities none of the time
4. You have pain that does not interfere with your normal work (both outside the home and housework) at all
5. You feel downhearted and low none of the time
6. You have a lot of energy most of the time
7. You have just had a consultation with the doctor. You evaluated your communication experience as very good.

The aim of this project is to explore respondents' reactions and thoughts to these alternative health states, which include both health and process outcomes.

iv. The project's methodology:

(this must be in language comprehensible to a lay person)

This project will use the 'think aloud' method and qualitative semi-structured interviews. The think aloud method is one whereby respondents are encouraged to speak and literally 'think aloud' as they solve a task. This allows the interviewer to record the thought process of the respondent concurrently as the tasks are being solved, rather than asking the participant to retrospectively remember what they did.

This is followed by a semi-structured interview. The interview allows us to explore the participant's thought process in-depth. The interview will probe further in depth on the criteria that the participants have mentioned. The interview will follow a topic guide.

Respondents will be presented with 4 health states, two of which will include only health outcomes, and two will include both health and process outcomes. They will be asked to value them using both the visual analogue scale, and time trade off techniques.

The VAS uses a rating scale from 0 to 100, where 0 is the "worst imaginable health state" and 100 is the "best imaginable health state". Respondents are asked to locate the health state descriptions on the scale between the two anchors so that the intervals between the scores correspond to the differences they perceive (Drummond et al., 2005).

The time trade off (TTO) will be based on the Measurement and Valuation of Health study (Dolan, 1997), which has provided the national values for the UK and has served as a guide for other national studies (Tsuchiya et al., 2002; Shaw et al., 2005; Lamers et al., 2006).

When the TTO technique is used to value temporary health states, 2 steps are completed: In step 1, a participant is asked to decide on a point of indifference between a certain time (t) spent in a temporary health state (h_i) followed by full health, and a lesser time spent (x) in a worse health state(h_j) followed by full health. Time x is varied until the participant reaches a point of indifference between the health states. The participant is again asked to decide between 2 choices: a certain time (t) in health state (h_j) followed by immediate death or a lesser time spent in full health followed by immediate death. These 2 steps, allow the calculation of the valuation for h_i and h_j.

To introduce respondents to the health state valuation task and the think aloud process, respondents will complete a practice task and asked to think aloud while performing the task. They will initially be asked a set of questions about themselves. They will then be asked to value the health states using the VAS and TTO.

This is followed by the semi-structured interview, which will begin with questions to explore the thoughts participants had whilst valuing the health states. For example, if they queried the relevance of a question, or appeared to struggle to answer it, the issue will be explored in the semi-structured interview.

The interview will then be based on a topic schedule which includes the following questions:

1. How did you find the think aloud exercise?
2. Were there questions that you found more difficult to answer than others?
3. Were things important to you during the exercise?
4. What did you think about the health states themselves?
5. Did the VAS and TTO make you think about things differently?
6. Did this alter when communication was added to the health state description?

The full topic schedule can be seen in the questionnaire booklet.

Analysis of the data will follow the Framework thematic method (Ritchie et al., 1994). The framework method is used regularly and is selected for its rigorous and transparent approach to the analysis of qualitative data. This method consists of a five step procedure: (1) familiarisation, (2) identifying themes, (3) indexing, (4) charting, (5) mapping and interpretation.

A6. What is the potential for physical and/or psychological harm/distress to participants?

The project may cause some inconvenience to participants, as they need to obtain transport to, and spend time at the university campus. However, it is good practice to conduct 'think aloud' interviews in a quiet setting (Fonteyn et al., 1993), and the university campus may be more appropriate than the

participant's home.

There is low potential for physical harm, as participants are not required to engage in any physical activity during the interview. The participants will be interviewed on campus and no physical harm is foreseen.

Participants will be required to value health states and as such may need to think about bad health and may remember or imagine experiences with bad health. In cases of severe health states they may need to consider whether a health state is worse than being dead. This may cause discomfort for participants. However, similar interviews have been conducted (Robinson et al., 1997; Baker and Robinson., 2004; Damschroder et al., 2005; Osch et al., 2007) and no harm is reported. The interview will be terminated if the interviewer feels the task is causing emotional distress.

A7. Does your research raise any issues of personal safety for you or other researchers involved in the project? (especially if taking place outside working hours or off University premises)

The aim is for all interviews to be conducted on university premises (White Rose room or the HSR Library in ScHARR) which should provide a safe and secure environment. The aim is to conduct interviews during working hours, which should minimize any potential harm. Only if it is not feasible to recruit enough participants will interviews be conducted off university premises during working hours.

If yes, explain how these issues will be managed.

Not applicable.

A8. How will the potential participants in the project be:

iv. **Identified?** *(please ensure that all practical issues about contacting individuals are covered and that you are not requesting the personal details of individuals be given over without their consent)*

The project requires people recruited from the 'general public', as is often done for studies obtaining values for health outcomes. No participants who are underage or who lack mental capacity will need to be recruited.

Participants will be selected to cover a broad age range and both genders. We aim to initially recruit 20 people, with the potential to recruit an additional 10. All participants will be recruited from Sheffield.

The recruitment target is set based on i) the qualitative nature of the study (and the potential point of theoretical saturation), ii) resource constraints, iii) the sample size from a previous and similar PhD study performed at Sheffield University which included 20 participants before reaching a point of theoretical saturation (whereby no new information was being retrieved) and iv) the findings from similar published studies (Robinson et al., 1997, 43 participants; Baker and Robinson 2004, 28 participants; Damschroder et al., 2005, 64 participants; Osch et al., 2007, 45 participants).

v. Approached?

Potential participants will be approached via the University of Sheffield mailing list, using the “announce” facility. This facility will be used to target non-academic staff, not working within the Economics department or ScHARR. Those expressing an interest to participate will be asked to respond with details such as their age and gender; as discussed above, the study sample will then be selected to be representative of the general population, based on census data reporting age and gender distributions for the UK. Those not selected will be sent a response thanking them for their interest; those selected will be emailed an information leaflet, including VB’s contact details for arranging an interview time. All interviews will be undertaken within Sheffield University. Those who are approached to participate in the study will be provided with an information sheet and consent form. If they express an interest in the study, an interview will be scheduled and consent obtained by the lead researcher (VB).

vi. Recruited?

People who are approached will be given the information sheet and consent form. If they express interest, an interview will be scheduled and consent obtained. The interview will be scheduled at least one week later to allow the participants’ time to consider the study and ask more questions. The consent form will be reviewed before the interview starts, where the participant may withdraw consent.

A9. Will informed consent be obtained from the participants?

Yes	X	No	

If informed consent or consent is NOT to be obtained please explain why. Further guidance is at: www.shef.ac.uk/ris/other/gov-ethics/ethicspolicy/policy-notes/consent

Not applicable.

A9.1. This question is only applicable if you are planning to obtain informed consent:

How do you plan to obtain informed consent? (i.e. the proposed process?):

Written consent forms will be distributed to people who express interest. Participants are asked to sign the consent form if they want to take part in the study. Before each interview, the interviewer will ask the participant if they have any more questions. They will be reminded that they may withdraw consent at anytime without providing any reasons.

Remember to attach your consent form and information sheet (where appropriate)

A10. What measures will be put in place to ensure confidentiality of personal data, where appropriate?

(As a minimum please ensure details are included of: how long data will be kept; when and how it will be destroyed; that PCs and other devices are password protected; that personal details are encrypted. This information should also be included on your information sheet).

Data will consist of notes made by the interviewer and audio recordings. All information will be anonymised. The audio recordings and notes will be identified using a respondent number. No other identifiable data will be included. The PhD student will transcribe the notes into Microsoft Word documents after each interview. The handwritten notes will be shredded after transcription.

The PhD student will also transcribe the audio-recordings. These transcripts will then be anonymised, removing all personal and identifying references. The transcripts and the audio will be saved on an encrypted and password-protected external hard-drive.

A11. Will financial/in kind payments (other than reasonable expenses and compensation for time) be offered to participants? (Indicate how much and on what basis this has been decided)

All interviewees will receive a £10 voucher for participating. This amount was based on previous studies completed within ScHARR (Mulhern et al., 2012). The funding for these incentives has been agreed by Professor Simon Dixon.

A12. Will the research involve the production of recorded media such as audio and/or video recordings?

Yes [X] No [ ]

A12.1. This question is only applicable if you are planning to produce recorded media:

How will you ensure that there is a clear agreement with participants as to how these recorded media may be stored, used and (if appropriate) destroyed?

The information and consent forms will inform participants that the interview is being audio recorded and how the data will be stored and used. Before each interview the participant will be asked if they consent to the recording being made and only then will the device be turned on.

The information form will make clear where the data is to be stored and who has access to this data, as well as the time the data is destroyed. It will also mention that anonymised data will be used to write up the PhD thesis and may potentially be used for conference presentations or publications. This will be summarized to the participants before the interview, stressing the fact

that we will collect audio recordings, anonymised transcripts of the audio recordings and field notes. These will be accessible only by the project team (PhD student and two supervisors). Analysis of their answers will also be used, in anonymised format, in the study write-up.

Guidance on a range of ethical issues, including safety and well-being, consent and anonymity, confidentiality and data protection are available at: www.shef.ac.uk/ris/other/gov-ethics/ethicspolicy/policy-notes

University Research Ethics Application Form - Part B - The Signed Declaration

Title of Research Project:

A "think aloud" study: Investigating how people respond to SF-7D health states.

I confirm my responsibility to deliver the research project in accordance with the University of Sheffield's policies and procedures, which include the University's '*Financial Regulations*', '*Good Research Practice Standards'* and the '*Ethics Policy Governing Research Involving Human Participants, Personal Data and Human Tissue'* (Ethics Policy) and, where externally funded, with the terms and conditions of the research funder.

In signing this research ethics application form I am also confirming that:

- The form is accurate to the best of my knowledge and belief.
- The project will abide by the University's Ethics Policy.
- There is no potential material interest that may, or may appear to, impair the independence and objectivity of researchers conducting this project.

- Subject to the research being approved, I undertake to adhere to the project protocol without unagreed deviation and to comply with any conditions set out in the letter from the University ethics reviewers notifying me of this.
- I undertake to inform the ethics reviewers of significant changes to the protocol (by contacting my academic department's Ethics Administrator in the first instance).
- I am aware of my responsibility to be up to date and comply with the requirements of the law and relevant guidelines relating to security and confidentiality of personal data, including the need to register when necessary with the appropriate Data Protection Officer (within the University the Data Protection Officer is based in CiCS).
- I understand that the project, including research records and data, may be subject to inspection for audit purposes, if required in future.
- I understand that personal data about me as a researcher in this form will be held by those involved in the ethics review procedure (e.g. the Ethics Administrator and/or ethics reviewers) and that this will be managed according to Data Protection Act principles.
- If this is an application for a 'generic' project, all the individual projects that fit under the generic project are compatible with this application.

- **I understand that this project cannot be submitted for ethics approval in more than one department, and that if I wish to appeal against the decision made, this must be done through the original department.**

Name of the Supervisor:

Dr Georgina Jones

Name of the student:

Victoria Brennan

Signature of the Supervisor:

[redacted]	**Date: 12th February 2014**

Email the completed application form and provide a signed, hard copy of 'Part B' to the Ethics Administrator (also enclose, if relevant, other documents).

References

Baker R, Robinson A Responses to standard gambles: are preferences 'well constructed'? Health Economics 2004 Jan 13(1):37-48.

Damschroder, osovic, M.E. & Ubel, P., 2005. Trading people ver e? Population health metrics, 3(10)

Dolan, P. (199 ealth States. Medical care, 35(11): 1095–1108.

Drummond MF, Sculpher MJ, Torrance GW, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes. Third Edition, Oxford. 2005.

Lamers, L M, McDonnell, J, Stalmeier, P F M, Krabbe, Paul and Busschbach, J J V (2006), The Dutch Tariff: Results and Arguments for an Effective Design for National EQ-5D Valuation Studies. Health economics, 15(10): 1121–32.

Mulhern B, Tsuchiya A, Brazier J, Rowen D. How do respondents perceive health state valuation exercises? A 'think aloud' study investigating time trade off and discrete choice experiments. Presented at EuroQol Plenary 2012 (Rotterdam).

Osch, S. van & Stiggelbout, A., 2008. The construction of standard gamble utilities. Health economics, 40(April 2008), pp.31–40.

Ritchie, J, and Lewis, J. (eds) (2003) Qualitative Research Practice Guide for Social Science Students and Researchers, Sage: A London.

Robinson, A., Dolan, P. and Williams, A. (1997) Valuing health status using VAS and TTO: what lies behind the numbers? Social Science and Medicine 45 (8), 1289-1297.

Shaw, J. W., Johnson, J. and Coons, S. J. (2005), US Valuation of the EQ-5D Health States: Development and Testing of the D1 Valuation Model. Medical care, 43(3): 203–20.

Spencer, A. (2001) The time trade-off method: an exploratory study [working paper 437]. London: Queen Mary College, University of London.

Tsuchiya, A., Ikeda, S., Ikegami, N., Nishimura, S., Sakai, I., Fukuda, T., Hamashima, C.,

What is the project's purpose?

To improve decision-making about funding different health care programmes, health economists aim to measure and compare the benefits of different health care treatments. These benefits are often measured by asking the general public their views about the desirability of different health state descriptions.

The health state descriptions used in these exercises most commonly include only health outcomes, for example they may include the following information:

1. Your health does not limit you in moderate activities
2. You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems
3. Your health limits your social activities none of the time
4. You have pain that does not interfere with your normal work (both outside the home and housework) at all
5. You feel downhearted and low none of the time
6. You have a lot of energy most of the time

However, a recent study performed at Sheffield University also included information relating to the process of receiving health care:

1. Your health does not limit you in moderate activities
2. You have no problems with your work or other regular daily activities as a result of your physical health or any emotional problems.
3. Your health limits your social activities none of the time
4. You have pain that does not interfere with your normal work (both outside the home and housework) at all

5. You feel downhearted and low none of the time
6. You have a lot of energy most of the time
7. **You have just had a consultation with the doctor. You evaluated your communication experience as very good.**

The aim of this project is to explore respondents' reactions and thoughts to these alternative health states, which include both health and process outcomes.

Why have I been chosen?

We aim to interview twenty to thirty members of the general public for this project.

Do I have to take part?

It is up to you to decide whether or not to take part. If you do decide to take part, you will be given a consent form to sign. You can still withdraw at anytime, you do not have to give a reason and there will be no penalty. No more information will be collected if you decide to withdraw from the study.

What do I have to do?

You will be invited at your convenience to the University of Sheffield campus. We will conduct an interview that should last about one hour. We will ask your opinion on different health state descriptions to 1) see which one you value the most, and 2) obtain your thoughts on the health states including both health and process outcomes.

Will I be recorded, and how will the recorded media be used?

We will make audio recordings during the interview and take notes during the interview. The audio recordings made during the interview will be transcribed and all personal references removed. No one outside the project team will be allowed access to the original recordings. The recordings will be destroyed after the project is over (in approximately two years time).

What are the possible benefits of taking part?

This project will help researchers better understand what aspects of health matter for people.

if something goes wrong?

In the first instance you may contact the PhD student (contact information is listed at the end). You may also contact the supervisors for this research, Dr Georgina Jones or Professor Simon Dixon, at:

The School of Health and Related Research (ScHARR)

Section of Health Economics & Decision Science

The University of Sheffield
Regent Court, 30 Regent Street
Sheffield, S1 4DA

Will my taking part in this project be kept confidential?

The information that we collect will be kept strictly confidential. You will not be identified or identifiable in any reports or publications. The audio recordings will be

transcribed and all references to personal details removed. The researchers will use the anonymised transcripts, not the actual voice recordings. After the project is finished the audio recordings will be destroyed.

What will happen to the results of the research project?

The results of this project will be used for a subsequent research project and form part of a PhD thesis. Some of the data may be used for journal articles or conference presentations. All results that are made public will be anonymised, by removing any words that could identify you, and you will not be identifiable in any publications.

Who is organising and funding the research?

The research is being organised by The University of Sheffield, as part of a PhD study that is funded by the PhD Students employer, RTI-Health Solutions.

Who has ethically reviewed the project?

The School of Health and Related Research of the University of Sheffield has reviewed and approved this project.

Contact for further information

You may contact Victoria Brennan

Telephone: 07725010182

Email: vkbrennan@hormail.com

Thank you for considering this research project.

Appendix 30: Think Aloud ScHARR Ethics Committee Approval

The University Of Sheffield.

Kirsty Woodhead
Ethics Committee Administrator

Regent Court
30 Regent Street
Sheffield S1 4DA
Telephone: +44 (0) 114 2225453
Fax: +44 (0) 114 272 4095 (non confidential)
Email: k.woodhead@sheffield.ac.uk

Our ref: 0715/KW

14 February 2014

Victoria Brennan
ScHARR

Dear Victoria

A "think aloud" study: Investigating how people respond to SF-7D health states.

Thank you for submitting the above research project for approval by the ScHARR Research Ethics Committee. On behalf of the University Chair of Ethics who reviewed your project, I am pleased to inform you that on 01 April 2014 the project was approved on ethics grounds, on the basis that you will adhere to the documents that you submitted for ethics review.

The research must be conducted within the requirements of the hosting/employing organisation or the organisation where the research is being undertaken. You are also required to ensure that you meet any research ethics and governance requirements in the country in which you are researching. It is your responsibility to find out what these are.

If during the course of the project you need to deviate significantly from the documents you submitted for review, please inform me since written approval will be required. Please also inform me should you decide to terminate the project prematurely.

Yours sincerely

[redacted]

Kirsty Woodhead
Ethics Committee Administrator

Appendix 31: Think Aloud Recruitment Email

I am PhD student working on a study to explore the thought processes people go through when they are asked to value different descriptions of health.

We are looking for between 20 and 30 volunteers to complete a face to face interview which will be held at Sheffield University, and will last between 45 minutes to 1.5 hours. You will be shown descriptions of health states and be asked to think out loud while you value them.

You will be reimbursed with a 10 pound gift voucher for your time.

Participants are not limited to university staff or students. Please feel free to pass this information on to anyone who might be interested.

If you are interested in taking part, we will provide you with a list of scheduled meeting dates and times to choose from. If you are interested or have any questions, please feel free to contact Victoria Brennan atvkbrennan@hotmail.com.

A copy of the study information sheet can be seen here:

https://www.dropbox.com/s/3l1jos8f0ly2xlb/Think%20Aloud%20Information%20sheet%20Final.doc (This needs entering into the web address field for it to be active).

The ScHARR ethics committee has approved this study. This PhD study is supervised by Simon Dixon and Georgina Jones (School of Health and Related Research).

Appendix 32: Think Aloud Framework Matrices

Table 66. Framework matrix theme 1 (Sections 1.1 and 1.2).

	1.1 Don't understand what it means.	1.2 Interpretation of bad process domain						
		1.2.1 Lack of support /help from doctor	1.2.2 Reflection of communication			1.2.3 Reflection of health	1.2.4 Other	
Transcript			1.2.2.1 Doctor and patients communication	1.2.2.2 Doctors communication	1.2.2.3 Patients communication		1.2.4.1 Not being listened to	1.2.4.2 Reflection of outcome
F1		And you've got a crappy doctor.						
F2		if your doctor is not on your side				but you've evaluated that you're feeling like this,	There is nothing worse than going to the doctors and just (sighs), and just being not listened to Hmmmm, not he's not listening to either of the times.	
F3							so you might have a, a bad experience going to the doctors and feel that you're	

	1.1 Don't understand what it means.	1.2 Interpretation of bad process domain						
		1.2.1 Lack of support /help from doctor	1.2.2 Reflection of communication			1.2.3 Reflection of health	1.2.4 Other	
Transcript			1.2.2.1 Doctor and patients communication	1.2.2.2 Doctors communication	1.2.2.3 Patients communication		1.2.4.1 Not being listened to	1.2.4.2 Reflection of outcome
							not being listened to ,	
F4					The doctor will only know what you reveal as the patient.			
F5			You've tried to explain to your doctor how you're feeling and um, what, I'm assuming, or maybe I'm assuming that, um, you haven't communicated very well, but it could be that the doctor possibly hasn't received what you are saying very well. being able to have good communication both ways with the doctor,		So you might come away thinking I wish I'd said that, or I wish I'd asked that, or whatever, but this one is really poor.			
F6	My communication experience is poor. Like with the doctor. Or?		I feel like I wasn't good enough at saying telling what I feel like, is that right? And that the doctor didn't give the proper feedback maybe?			OK, so the communication experience is poor and the health state is also poorer. The perceived health		

	1.1 Don't understand what it means.	1.2 Interpretation of bad process domain						
		1.2.1 Lack of support /help from doctor	1.2.2 Reflection of communication			1.2.3 Reflection of health	1.2.4 Other	
Transcript			1.2.2.1 Doctor and patients communication	1.2.2.2 Doctors communication	1.2.2.3 Patients communication		1.2.4.1 Not being listened to	1.2.4.2 Reflection of outcome
						state I guess.		
F7						I think the last thing that's key here is the last bullet point, about communicating with your doctor, your health status is poor, because I mean, you actively saying that and actively feeling that, it takes a lot to admit that you're not in the best health possible. So its quite similar to the other ones, but here, the fact that you've actually communicated that, that makes you feel quite		

	1.1 Don't understand what it means.	1.2 Interpretation of bad process domain						
		1.2.1 Lack of support /help from doctor	1.2.2 Reflection of communication			1.2.3 Reflection of health	1.2.4 Other	
Transcript			1.2.2.1 Doctor and patients communication	1.2.2.2 Doctors communication	1.2.2.3 Patients communication		1.2.4.1 Not being listened to	1.2.4.2 Reflection of outcome
						poor, um in terms of your health status, I'm going to say, um, 40. Because you've communicated that all of those things are going on.		
F8								
F9								I think as I was going along I was sort of jumping to the top and then jumping to the bottom to see what the outcome was, a little bit more for the doctor.
M1		Whereas this person looks like they are not getting the help they need.						

	1.1 Don't understand what it means.	1.2 Interpretation of bad process domain						
		1.2.1 Lack of support /help from doctor	1.2.2 Reflection of communication			1.2.3 Reflection of health	1.2.4 Other	
Transcript			1.2.2.1 Doctor and patients communication	1.2.2.2 Doctors communication	1.2.2.3 Patients communication		1.2.4.1 Not being listened to	1.2.4.2 Reflection of outcome
M2	Can I just ask when it says communication experience, what does that really, what does that mean?		Not having had a good communication good with your doctor whether it's how you have portrayed it or he has received it.					
M3	Can I just ask about the um, consultation thing, um, when, when its evaluated as communicated as communication experience as good, does that mean good news was given, or just kind of like you enjoyed having a chat.							
M4								
M5								
M6								
M7	I'm not sure what conclusions to draw from that. I was a bit puzzled by that really.		I'm assuming from the question that it's possibly the health that um prevents you from communicating properly with the doctor. And then it can be that the doctor			I'm assuming from the question that it's possibly the health that um prevents you from communicating properly with the		

	1.1 Don't understand what it means.	1.2 Interpretation of bad process domain						
		1.2.1 Lack of support /help from doctor	1.2.2 Reflection of communication			1.2.3 Reflection of health	1.2.4 Other	
Transcript			1.2.2.1 Doctor and patients communication	1.2.2.2 Doctors communication	1.2.2.3 Patients communication		1.2.4.1 Not being listened to	1.2.4.2 Reflection of outcome
	it wasn't too sure what conclusions to draw from that really.		isn't communicating very well.			doctor.		
M8	In here it says communication experience. What does that mean?					he decided to tell me that I've got some problem in me then obviously you know I may not be as good as I think. and the doctor say I am poorly. the doctor say like communication experience is poor, so he suggests that you know, I have problems.		
M9	you have just consulted with your doctor, your communication experience was poor. So the doctor thinks your communication							

	1.1 Don't understand what it means.	1.2 Interpretation of bad process domain						
		1.2.1 Lack of support /help from doctor	1.2.2 Reflection of communication			1.2.3 Reflection of health	1.2.4 Other	
Transcript			1.2.2.1 Doctor and patients communication	1.2.2.2 Doctors communication	1.2.2.3 Patients communication		1.2.4.1 Not being listened to	1.2.4.2 Reflection of outcome
	experience is poor? So, like, he wasn't a good doctor kind of like. You didn't think like, he kind of helped you, or?							
M10	Um, can I just ask what it means by communication experience?	don't feel like your doctors giving you much support,						
M11								

Table 67. Framework matrix theme 1 (Sections 1.3).

	1.3 Interpretation of a good process domain								
	1.3.1 Being listened to	1.3.2 Doctor helps	1.3.3 Reflection of communication			1.3.4 Reflection of health	1.3.5 Reflection of hope	1.3.6 Reflection of outcome/ progress	1.3.7 Other
Transcript			1.3.3.1 Doctor and patients communication	1.3.3.2 Doctors communication	1.3.3.3 Patients communication				
F1							if I've had a good experience, I feel like it's on the way to being sorted.	If I go to the doctors for a health issue even just seeing the doctor if I've had a good experience, I feel like it's on the way to being sorted.	
F2	yep, the same scenario again, um, but if you, if you have been to the doctors, and they've said, oh yes, it's your thyroid and we can do something about that, and this is why you're tired all of the time, and this is	oh hang on, sorry. Bloody doctors, so. Um, (sniffs) the first one, Q10 is where you would get help from your GP					I would definitely take the 4 weeks, (giggles) with hope, um, uhh, and you know the doctor saying we can sort this out, I can send you for some tests. You take these antibiotics. Whenever, you know, this is		

	1.3 Interpretation of a good process domain								
	1.3.1 Being listened to	1.3.2 Doctor helps	1.3.3 Reflection of communication			1.3.4 Reflection of health	1.3.5 Reflection of hope	1.3.6 Reflection of outcome/ progress	1.3.7 Other
Transcript			1.3.3.1 Doctor and patients communication	1.3.3.2 Doctors communication	1.3.3.3 Patients communication				
	why you don't feel so well sometimes, and he's listened to you and understood you, then you feel that there more hope, and it's just, it's not you						something we can sort out. I would much rather have 4 weeks with, with the doctors support and the hope of getting better than any time with just getting foistered off, um coz there is nothing worse than going to the doctors and just (sighs), and just being not listened to, so um.		
F3									
F4									

	1.3 Interpretation of a good process domain								
	1.3.1 Being listened to	1.3.2 Doctor helps	1.3.3 Reflection of communication			1.3.4 Reflection of health	1.3.5 Reflection of hope	1.3.6 Reflection of outcome/ progress	1.3.7 Other
Transcript			1.3.3.1 Doctor and patients communication	1.3.3.2 Doctors communication	1.3.3.3 Patients communication				
F5					Although you've got a lot of the same things going on, you can still explain to the doctor what's happening, I know myself, if do go to the doctor, and I feel I've been able to express what I want to say well, and um, I have a fantastic GP, and I just always feel better (giggles) when I've seen him.				
F6									
F7					But it says you communicated that your health status is good.				
F8									

	1.3 Interpretation of a good process domain								
	1.3.1 Being listened to	1.3.2 Doctor helps	1.3.3 Reflection of communication			1.3.4 Reflection of health	1.3.5 Reflection of hope	1.3.6 Reflection of outcome/ progress	1.3.7 Other
Transcript			1.3.3.1 Doctor and patients communication	1.3.3.2 Doctors communication	1.3.3.3 Patients communication				
F9								I think as I was going along I was sort of jumping to the top and then jumping to the bottom to see what the outcome was, a little bit more for the doctor.	
M1								And you had a good experience with the doctor, so things are probably on their way.	
M2							You place a lot of um value in what he or she has to say, and um, I would hope, it's the word I used before, it would	You know there is a chance that you can get better.	

	1.3 Interpretation of a good process domain								
	1.3.1 Being listened to	1.3.2 Doctor helps	1.3.3 Reflection of communication			1.3.4 Reflection of health	1.3.5 Reflection of hope	1.3.6 Reflection of outcome/ progress	1.3.7 Other
Transcript			1.3.3.1 Doctor and patients communication	1.3.3.2 Doctors communication	1.3.3.3 Patients communication				
							give you some hope and sort of um a better feeling. And that it's not um, (slaps lips) how to describe it. You know there is a chance that you can get better.		
M3								I mean I guess some people would be looking forward to having a good outcome from that. Sorry a good evaluation, coming to a good conclusion.	
M4									

	1.3 Interpretation of a good process domain								
	1.3.1 Being listened to	1.3.2 Doctor helps	1.3.3 Reflection of communication			1.3.4 Reflection of health	1.3.5 Reflection of hope	1.3.6 Reflection of outcome/ progress	1.3.7 Other
Transcript			1.3.3.1 Doctor and patients communication	1.3.3.2 Doctors communication	1.3.3.3 Patients communication				
M5									
M6									
M7				It could potentially be a doctor, very, you know well communicated things properly,	possibly the health state has prevented me from communicating with the doctor very well.				
M8									
M9									
M10	communication experience good, um I mean yes, so I suppose thinking about how I've felt when I've come out of the doctors, and the experience has been, you feel like the doctor has listened to you, and understands			you feel like the doctor has listened to you, and understands what you are going through, um.					Again, um, I feel that quality of life would be better um, for 2 years, then health state A for 5 years, especially, there's um, you'd feel more rubbish, um, if you didn't feel the doctor was supportive, um, so health state um, A

	1.3 Interpretation of a good process domain								
	1.3.1 Being listened to	1.3.2 Doctor helps	1.3.3 Reflection of communication			1.3.4 Reflection of health	1.3.5 Reflection of hope	1.3.6 Reflection of outcome/ progress	1.3.7 Other
Transcript			1.3.3.1 Doctor and patients communication	1.3.3.2 Doctors communication	1.3.3.3 Patients communication				
	what you are going through, um.								for 5 years versus 1.5 years, so, I think full health for 1.5 would be better again because of quality of life, um.
M11									

Table 68. Framework matrix theme 2 (Section 2.1).

	2.1 Impact of having a positive consultation.				
	2.1.1 Positive impact				
Transcript	**2.1.1.1 On the way you feel**	**2.1.1.2 Give hope or encouragement**	**2.1.1.3 On outlook**	**2.1.1.4 On health**	**2.1.1.5 Other**
F1	If I go to the doctors for a health issue even just seeing the doctor if I've had a good experience, I feel like it's on the way to being sorted.	You've just had a consultation with your doctor and you evaluated your communication experience as good. I think that persons going to have had a probably I mean, if I go to the doctors for a health issue even just seeing the doctor if I've had a good experience, I feel like it's on the way to being sorted.	You've just had a consultation with your doctor and you evaluated your communication experience as good. I think that persons going to have had a probably I mean, if I go to the doctors for a health issue even just seeing the doctor if I've had a good experience, I feel like it's on the way to being sorted. They might have been referred on for some sort of treatment towards what's wrong with them so I would rate that more highly, 85 probably, because you tend to have a more positive outlook on your health state if you've seen a doctor and have a good experience than if not you feel like it's sort of being treated as if it's a, whoops, sorry dropped it on the floor.		
F2	(giggles). So I think, just because you, that hope, and that sort of like, that something can be done about it, instead of being in like a fog where there is no way out. That instantly I think makes you feel a bit more positive about your health state. Because you know that something can be done	(giggles). So I think, just because you, that hope, and that sort of like, that something can be done about it, instead of being in like a fog where there is no way out.	(giggles). So I think, just because you, that hope, and that sort of like, that something can be done about it, instead of being in like a fog where there is no way out.	I can send you for some tests. You take these antibiotics.	

	2.1 Impact of having a positive consultation.				
	2.1.1 Positive impact				
Transcript	**2.1.1.1 On the way you feel**	**2.1.1.2 Give hope or encouragement**	**2.1.1.3 On outlook**	**2.1.1.4 On health**	**2.1.1.5 Other**
	about it.	That instantly I think makes you feel a bit more positive about your health state. Because you know that something can be done about it. Consultation with the doctor, and the experience is good, and then health state B, energy none of the time, downhearted and low all of the time, and then, no communication with the doctor. Yep. I would take, I would definitely take the 4 weeks, (giggles) with hope, um, uhh, and you know the doctor saying we can sort this out, I can send you for some tests. You take these antibiotics. Whenever, you know, this is something we can sort out. I would much rather have 4 weeks with, with the doctors support and the hope of getting better than any time with just getting foistered off, um coz there is nothing worse than going to the doctors and just (sighs), and just being not listened to, so um.	That instantly I think makes you feel a bit more positive about your health state. Because you know that something can be done about it.		

	2.1 Impact of having a positive consultation.				
	2.1.1 Positive impact				
Transcript	**2.1.1.1 On the way you feel**	**2.1.1.2 Give hope or encouragement**	**2.1.1.3 On outlook**	**2.1.1.4 On health**	**2.1.1.5 Other**
		I think if you go and there is a negative experience then you think, well I've got nowhere to go now. So it's that hope, a little bit,			
F3	this seems quite similar, um, but I guess having a positive experience with the consultation with your doctor would make you feel better about the situation, um, I think um that (slaps lips), about there. Ye, because it might just, um, reaffirm that you're in quite good health, and you're feeling quite, quite well, and the doctor might agree with you, and praise you. Ye.				
F4					You've still got emotional problems, but the doctor says the communication experience is good. But then does the doctor really know what's going on inside your head? Um. I mean for me, it's, it doesn't sound much different, apart from what the doctor is saying. Then I struggle to think, from an emotional point of view your doctor would really know what's going on if you've got emotional problems. The doctor will only know what you reveal as the patient. (Sighs),Hmmm. (Sighs) well that patient is communicating better. Um. I'd probably put that at maybe 45 then. Trusting that the patients being honest with the

	2.1 Impact of having a positive consultation.				
	2.1.1 Positive impact				
Transcript	**2.1.1.1 On the way you feel**	**2.1.1.2 Give hope or encouragement**	**2.1.1.3 On outlook**	**2.1.1.4 On health**	**2.1.1.5 Other**
					doctor. Because I know full well with my partner that he does not like doing to the doctor, and keeps things hidden.
F5	I think, um, the way we are treated by doctors can have a really positive or negative effect on, on health outcomes, and on the way you feel about, um, I mean it's not really saying that its symptoms you're making then um, but um, you can feel better can't you, if you've had a good visit to the doctor. Um, well I think I've just said something about that haven't I. That um, being able to have good communication both ways with the doctor, um, does have an um, a profound effect on how you feel about um, your health state. I would find that quite difficult to feel that			I think, um, the way we are treated by doctors can have a really positive or negative effect on, on health outcomes, and on the way you feel about, um, I mean it's not really saying that its symptoms you're making then um, but um, you can feel better can't you, if you've had a good visit to the doctor.	I know myself, if do go to the doctor, and I feel I've been able to express what I want to say well, and um, I have a fantastic GP, and I just always feel better (giggles) when I've seen him.
F6					
F7					
F8			This question appears to be the same as the other question apart from that the consultation with the doctor was good. So I'll put this back up to 80 because it would make you feel more positive about your outcomes and how things would progress.		
F9					
M1					And you had a good experience with the doctor, so things are

	2.1 Impact of having a positive consultation.				
	2.1.1 Positive impact				
Transcript	**2.1.1.1 On the way you feel**	**2.1.1.2 Give hope or encouragement**	**2.1.1.3 On outlook**	**2.1.1.4 On health**	**2.1.1.5 Other**
					probably on their way. I think the poor consultation with the doctor, just adds on top of that for this person. This is just adding on top of the rest of this. Whereas, because you've still got some kind of social activity here. And you still have lots of energy, you might have a poor consultation, but I don't think it's going to be as damaging in this case as it might be in this case, especially as its very poor and this is just poor. Not that's that's good either But I guess, it is more significant if you are feeling tense already and then you go to the doctor and don't get anything from that, and it is a very poor consultation. Then it adds there. But if you are feeling okish and you have a poor consultation. I still think you would get over it.
M2		You place a lot of um value in what he or she has to say, and um, I would hope, it's the word I used before, it would give you some hope and sort of um a better feeling.	You know there is a chance that you can get better.		
M3					
M4	And I would meet the doctor, and experience a				

	2.1 Impact of having a positive consultation.				
	2.1.1 Positive impact				
Transcript	**2.1.1.1 On the way you feel**	**2.1.1.2 Give hope or encouragement**	**2.1.1.3 On outlook**	**2.1.1.4 On health**	**2.1.1.5 Other**
	good. Yes. I would imagine if it was a long term issue then so that would be a push from time to time, but um, it depends if it um, long term situation that would make me happy, although I would respect, and I know So I um, that that would be the best the doctor could do, for example.				
M5					
M6		It gives you hope.			Oh, so this is good. So if you can have faith in your doctor and can talk through your problems then ye, that's probably, I'd probably rate it at 45, no 46, because it goes in scales of 2. I think if you have an understanding doctor, who has diagnosed you properly and has the experience of dealing with these kind of patients in the past it does help
M7					
M8			Um, I would say 50 in the last one, because you know the doctor say um, and communication experience is good. So it may not be as bad as I think.		
M9	OK, well this time you put the communication experience as good. So, you know, like I said it can change your perceptions and think that, you know, because it was good then ye, your health state can be a bit better I guess.				
M10	communication experience good, um I mean yes,				

	2.1 Impact of having a positive consultation.				
	2.1.1 Positive impact				
Transcript	**2.1.1.1 On the way you feel**	**2.1.1.2 Give hope or encouragement**	**2.1.1.3 On outlook**	**2.1.1.4 On health**	**2.1.1.5 Other**
	so I suppose thinking about how I've felt when I've come out of the doctors, and the experience has been, you feel like the doctor has listened to you, and understands what you are going through, um.				
M11	Ye, if it was poor then, I would feel less healthy but not as bad as the last time because its I mean as a practical person I'd just let it be he' take care of my health, but ye, I wasn't as good, if he encouraged me a little bit I might feel better so I'll put it at 55.	You have lot of energy some of the time, it is the same but you value your communication as good. So I think the communication can up bring up my hope a little so I'll out it at 55, the same as this one, even though my physical health is worse, but probably the doctor will encourage me a bit more.			

Table 69. Framework matrix theme 2 (Section 2.2).

	2.2 Impact of having a negative consultation.							
	2.2.1 Negative impact					2.2.2 Second opinion	2.2.3 Other	2.2.4 No impact
Transcript	**2.2.1.1 On the way you feel**	**2.2.1.2 Give hope or encouragement**	**2.2.1.3 On health**	**2.2.1.4 On coping with your illness**	**2.2.1.5 Other**			
F1	I think you'd be down regularly, I think that would make me sad.			Um, and the doctor, I think adds a huge influence on all of those, on, especially on the kind of mentality of a person when they go into a doctor if you're feeling low and your doctor picks you up that has an impact on how you approach your disease		you'd probably change doctor if you felt that strongly about it,		
F2	you've evaluated that you're feeling like this, and you've gone to your GP and if they've just fobbed you off, that makes you feel worse quite a lot of the time it's just going to be a bloody long old battle. if you go and there is a negative experience then you think, well I've got nowhere to go now. You assume that you're going	I think, me personally, if you've got pain or if you're feeling low, and um, you know your life's limited by things., or if you, you've always got that hope that if you go to your GP They'll always be able to magically sort it for you, or give you something. And I think that's just how most people view GPs, but I think if you go to your				my previous one where I could handle it for a week, um, and just change doctor or something. I probably wouldn't change doctor, but I'd um, I'd think. Think twice twice about going to him with something, something um, similar again, um,		

	2.2 Impact of having a negative consultation.							
	2.2.1 Negative impact					2.2.2 Second opinion	2.2.3 Other	2.2.4 No impact
Transcript	2.2.1.1 On the way you feel	2.2.1.2 Give hope or encouragement	2.2.1.3 On health	2.2.1.4 On coping with your illness	2.2.1.5 Other			
	to get sorted. You just do, and then when you realise you don't, then you go into the pits of despair. Or someone says you're just making it up in your mind, there's nowhere to go from there, um and if you haven't got the, the strength of mind because you're feeling downhearted and low, and you haven't got the friends to talk to, then you have a bad experience with your GP. Where are you going to go from there? Mentally?	GP and you get fobbed off, or um, they say, you know, just do this or do this, or exercise more, it's not that helpful to you, um. And I think a lot of the time you just want, I don't know, just to be, sort of listened to and have something that you can um, deal with. And work on, so um, ye, I think to be ye, that's not the best state, because of the, hmmm, ok, I would say about just because you feel you've got nowhere to go after that and you'd have to look into alternative therapies which is just time consuming (giggles) I would say about, fifty, no about six, no, 58 just because of the lack of hope. I think if you go and there is a negative experience then you				but I think again I mean you could change doctors obviously but if he's just going "pick yourself up" and you're going "I can't" and he's not offering you any alternatives, then it's just going to be a bloody long old battle. You could go to another doctor, there are quite a lot of people who you know what ever medical people say goes. For example, my dad, like he says well that's what the doctor says, so that must be true, and I say, well there are second opinions dad. But, so quite a lot of people are in the		

	2.2 Impact of having a negative consultation.							
	2.2.1 Negative impact					2.2.2 Second opinion	2.2.3 Other	2.2.4 No impact
Transcript	**2.2.1.1 On the way you feel**	**2.2.1.2 Give hope or encouragement**	**2.2.1.3 On health**	**2.2.1.4 On coping with your illness**	**2.2.1.5 Other**			
		think, well I've got nowhere to go now. So it's that hope, a little bit, For example, my dad, like he says well that's what the doctor says, so that must be true, and I say, well there are second opinions dad. But, so quite a lot of people are in the mind-set that the medical profession are always right. So wouldn't seek out a second opinions, and also, if they are of a low mood anyway. They wouldn't seek out that second opinions. So I think that's a massive, massive thing. Because then that hopes gone then for someone. it, alters you, it alters your view on hope.				mind-set that the medical profession are always right. So wouldn't seek out a second opinions, and also, if they are of a low mood anyway. They wouldn't seek out that second opinions. So I think that's a massive, massive thing.		
F3	where it says I had a consultation with the doctor and the communication							

	2.2 Impact of having a negative consultation.							
	2.2.1 Negative impact					2.2.2 Second opinion	2.2.3 Other	2.2.4 No impact
Transcript	**2.2.1.1 On the way you feel**	**2.2.1.2 Give hope or encouragement**	**2.2.1.3 On health**	**2.2.1.4 On coping with your illness**	**2.2.1.5 Other**			
	experience was poor, I think that would make me feel worse about the situation, even though even though, you know, my health looks the same as it was on the last question. But I think it's just, ye, that that would feel um, you might feel quite upset about it, I was picturing myself sort of at my doctors surgery and how I would feel leaving the building, and how I might feel very angry. Thinking how angry I would feel having that experience at a doctors.							
F4								
F5			just having had rubbish experiences with doctors, um, you do feel worse, um, and it depends on your kind of emotional mood as well, so it could have, if you are feeling rubbish, um, emotionally it's when you've had					

	2.2 Impact of having a negative consultation.							
	2.2.1 Negative impact					2.2.2 Second opinion	2.2.3 Other	2.2.4 No impact
Transcript	**2.2.1.1 On the way you feel**	**2.2.1.2 Give hope or encouragement**	**2.2.1.3 On health**	**2.2.1.4 On coping with your illness**	**2.2.1.5 Other**			
			other problems, then I know I've come out and it had impacted on the way I'd felt and perhaps hindered my um, kind of improvement in health,					
F6								
F7								
F8	um they've had a poor communication experience with the doctor. Um. So I'd value this a little bit lower, because that would make you feel a bit downhearted about this, and less positive about things. The consultation with the doctor, it didn't affect the quality of life, although it probably would have made you feel a bit low. I know it's disappointing to have a poor consultation with your doctor. But I think in the wider scheme of things I'm only interested in the actual health.							If you have a good consultation with the doctor it's not going to do anything to change your health.
F9					I've jumped to the			

	2.2 Impact of having a negative consultation.							
	2.2.1 Negative impact					2.2.2 Second opinion	2.2.3 Other	2.2.4 No impact
Transcript	**2.2.1.1 On the way you feel**	**2.2.1.2 Give hope or encouragement**	**2.2.1.3 On health**	**2.2.1.4 On coping with your illness**	**2.2.1.5 Other**			
					last line as well and its says your communication experience was very poor, so obviously it's not looking quite as good.			
M1					Um, I think the poor consultation with the doctor, just adds on top of that for this person. This is just adding on top of the rest of this.			
M2	And you feel unfulfilled, and their response is that you don't think that you have got what you were looking for. although the um, issues are lesser with A than they are with B, they are still severe, so having a negative experience with the doctor, or a non-productive experience with the doctor is um going to heighten that more for you, its not going to make you feel better	Not having had a good communication good with your doctor whether it's how you have portrayed it or he has received it sort of doesn't um, give you any hope or anything like that,		but as far as the doctor is concerned, there are alternates to the things which doctors can advise, um.				
M3						Perhaps I would try another doctor if I	I wouldn't think it was the end of the	

	2.2 Impact of having a negative consultation.							
	2.2.1 Negative impact					2.2.2 Second opinion	2.2.3 Other	2.2.4 No impact
Transcript	2.2.1.1 On the way you feel	2.2.1.2 Give hope or encouragement	2.2.1.3 On health	2.2.1.4 On coping with your illness	2.2.1.5 Other			
						wasn't happy or something like that I mean I wouldn't think it was the end of the world, that's for sure.	world, that's for sure.	
M4						Well my attitude, toward doctors like um, if I'm not satisfied with one, I would again try another one I mean I wouldn't go, I would just go to another one, another doctor. I wouldn't bother.		
M5								
M6								
M7						In that case you might just try to see another doctor to get a second opinion.		
M8								
M9					Well I guess, with the doctor, if the communication with the doctor was poor,			

	2.2 Impact of having a negative consultation.							
	2.2.1 Negative impact					2.2.2 Second opinion	2.2.3 Other	2.2.4 No impact
Transcript	2.2.1.1 On the way you feel	2.2.1.2 Give hope or encouragement	2.2.1.3 On health	2.2.1.4 On coping with your illness	2.2.1.5 Other			
					um, it makes it, it can have more of a kind of negative effect with kind of what you perceive of your health state I guess.			
M10	I suppose it's the, it's the same as the last one except you've just had a rubbish experience with your doctor, and how that's made me feel, or you feel, gloomier about your, the outcome or, gloomier about how you might deal with the problems you're facing so, um. Ye, that might not be a long term kind of impact on the health state, but it can have an um, impact, on initial moods,..... you'd feel more rubbish, um, if you didn't feel the doctor was supportive, and it depends on your kind of emotional mood as well, so it could have, if you are feeling rubbish, um, emotionally it's when you've had other problems, then I know I've come out and it had impacted	I suppose it's just kind of figuring out how I'd feel if I didn't think, that there wasn't much hope, um which I think with that consultation with the doctor, you might not feel that there's much room for improvement.					I suppose it's the, it's the same as the last one except you've just had a rubbish experience with your doctor, and how that's made me feel, or you feel, gloomier about your, the outcome or, gloomier about how you might deal with the problems you're facing so, um. Ye, that might not be a long term kind of impact on the health state, but it can have an um, impact, on initial moods, you'd feel more	I suppose you feel better than when you don't feel like um the doctors understood what's going on. Um, but then, I don't think overall, it would have that much impact.

	2.2 Impact of having a negative consultation.							
	2.2.1 Negative impact					2.2.2 Second opinion	2.2.3 Other	2.2.4 No impact
Transcript	**2.2.1.1 On the way you feel**	**2.2.1.2 Give hope or encouragement**	**2.2.1.3 On health**	**2.2.1.4 On coping with your illness**	**2.2.1.5 Other**			
	on the way I'd felt it might make you feel worse if your doctors kind of still, if you're still having poor experiences with them,						rubbish, um, if you didn't feel the doctor was supportive, and it depends on your kind of emotional mood as well, so it could have, if you are feeling rubbish, um, emotionally it's when you've had other problems, then I know I've come out and it had impacted on the way I'd felt it might make you feel worse if your doctors kind of still, if you're still having poor experiences with them,	
M11	Ye, if it was poor then, I would feel less healthy				So, if the doctor treat me very poorly,	Ye, if it was poor then, I would feel		

	2.2 Impact of having a negative consultation.							
	2.2.1 Negative impact					2.2.2 Second opinion	2.2.3 Other	2.2.4 No impact
Transcript	2.2.1.1 On the way you feel	2.2.1.2 Give hope or encouragement	2.2.1.3 On health	2.2.1.4 On coping with your illness	2.2.1.5 Other			
					would it affect my, I mean my decision to live? Um, no I wouldn't let him. I would still do what I want to do.	less healthy		

Table 70. Framework matrix theme 3.

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			**3.1.2.1.1 Improved health state valuation**	**3.1.2.1.2 No impact on health state valuation**	**3.1.2.1.3 Worsened health state valuation**	**3.1.2.2.1 Improved health state valuation**	**3.1.2.2.2 No impact on health state valuation**	**3.1.2.2.3 Worsened health state valuation**	
F1			They might have been referred on for some sort of treatment towards what's wrong with them so I would rate that more highly, 85 probably, because you tend to have a more positive outlook on your health state if you've seen a doctor and have a good experience than if not						
F2			(giggles). So I think, just because you, that hope,					I would say about, fifty, no about six, no, 58 just because of the lack of hope.	

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
			and that sort of like, that something can be done about it, instead of being in like a fog where there is no way out. That instantly I think makes you feel a bit more positive about your health state. Because you know that something can be done about it. So I'd say it would take you up to about, like 70 or 75 possibly just because you know it's something that you can work						

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
F3								I was going to place this at 80, because I think this looks quite similar wording, up to the one, um before the last one, um, but I guess just the last one where it says I had a consultation with the doctor and the communication experience was poor, I think that would make me feel worse about the situation, even though even though, you know, my health looks the same as it was on the last question. But I think it's just, ye, that that would feel um, you might feel quite upset about it, so I think I might rate it just below 80. And then, sometimes I think (sighs) you could talk to other people about your experience and that might make you feel better, so not just	

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
								the doctor, so you might have a, a bad experience going to the doctors and feel that you're not being listened to , but then you might be able to go home and talk it over with your friend of a partner, but I think I'd still maybe just um a little bit less than the first one, so I'd probably like about there, I think.	
F4								Hm, well the 2 things that popped out to me in there were this emotional problems, and communication experience is poor. My partner has emotional problems and it um, you know as a result of his depression and such, so it's um. And I know it can really limit his life a lot. So I'd say that was a little bit worse. And communicating, ye,	

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
								he's become quite agoraphobic and he sort of pulls back from society, so I would say that's a bit worse. I'd say its 40.	
F5		Right, I don't think this one is so difficult, um, I know it it's got a lot of the same things, but um, we haven't got the doctor bit, and um, tense and downhearted, but still, still only low a little bit of the time, so I think, I'll have it a little bit higher than the previous one maybe, go up to 40.						Um, you've just had a con… ye, I think that's quite, quite significant, that you've tried to explain to your doctor how you're feeling and um, what, I'm assuming, or maybe I'm assuming that, um, you haven't communicated very well, but it could be that the doctor possibly hasn't received what you are saying very well. So, I would find that quite difficult to feel that, so I think, I would put that at about 20. VB Yep V3 A1 OK (Reads through health	

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
								state silently) Right, I don't think this one is so difficult, um, I know it it's got a lot of the same things, but um, we haven't got the doctor bit, and um, tense and downhearted, but still, still only low a little bit of the time, so I think, I'll have it a little bit higher than the previous one maybe, go up to 40. What did I put there, it was 20, so 40. Because I would think not being able to explain to the doctor how you feeling is quite significant really.	
F6							I think the same, ye, 35, I don't think this will influence me.		
F7								So its quite similar to the other ones, but here, the fact that you've actually communicated that,	

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
								that makes you feel quite poor, um in terms of your health status, I'm going to say, um, 40. Because you've communicated that all of those things are going on.	
F8			This question appears to be the same as the other question apart from that the consultation with the doctor was good. So I'll put this back up to 80 because it would make you feel more positive about your outcomes and how things would progress.					Then this next question has got um a separate question at the bottom um although the health is limited only a little, um, um they've had a poor communication experience with the doctor. Um. So I'd value this a little bit lower, because that would make you feel a bit downhearted about this, and less positive about things. So I would rate this as 70.	

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
F9		Do you think it altered your value? The way you valued that health state? A7 Ye, definitely, um, because, I'm I think as I was going along I was sort of jumping to the top and then jumping to the bottom to see what the outcome was, a little bit more for the doctor. And thinking oh alright, this isn't going to be as well, or this is going to be better this one. So it sort of surmised it slightly.	So I think I'm going to give this one, um a slightly higher score, of nearly 60. Also the consultation with the doctor. It's been a good one as well. Do you think it altered your value? The way you valued that health state? A7 Ye, definitely, um, because, I'm I think as I was going along I was sort of jumping to the top and then jumping to the bottom to see what the outcome					Do you think it altered your value? The way you valued that health state? A7 Ye, definitley, um, because, I'm I think as I was going along I was sort of jumping to the top and then jumpting to the bottom to see what the oucome was, a little bit more for the doctor. And thinking oh alright, this isnt going to be as well, or this is going to be better this one. So it sort of summised it slightly.	

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			**3.1.2.1.1 Improved health state valuation**	**3.1.2.1.2 No impact on health state valuation**	**3.1.2.1.3 Worsened health state valuation**	**3.1.2.2.1 Improved health state valuation**	**3.1.2.2.2 No impact on health state valuation**	**3.1.2.2.3 Worsened health state valuation**	
			was, a little bit more for the doctor. And thinking oh alright, this isn't going to be as well, or this is going to be better this one. So it sort of surmised it slightly.						
M1									Reads question…… you have just had a consultation with your doctor, and you valued your communication experience as poor. Ye, so I still think that's, ye, ahhh, you are downhearted and low a little of the time. I'd still say that's pretty good. I'd still say its ok. I'd say 75.

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
M2		Right, ok, sure, um well, I'd I'd even place that shortly before on the scale, a little worse than the previous one, because the conditions are similar, but I think in terms of um, (lip slapping) your state of mind. Um. Not having had a good communication good with your doctor whether it's how you have portrayed it or he has received it sort of doesn't um, give you any hope or anything like that, so yes. OK, um, so it's the thing with the doctor. So that makes it worse. So that's quite easy. I'm just going to put crosses all of the way down. Do you think adding on the communication to that question made it easier or harder? P Um, well it made it easier for me. ye, it made it easier.							Right, ok, sure, um well, I'd I'd even place that shortly before on the scale, a little worse than the previous one, because the conditions are similar, but I think in terms of um, (lip slapping) your state of mind. Um. Not having had a good communication good with your doctor whether it's how you have portrayed it or he has received it sort of doesn't um, give you any hope or anything like that, so yes.	
M3				So I think the only difference					So I think the only difference is the

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
				is the um, consultation with the doctor. So, for me it's very similar,					um, consultation with the doctor. So, for me it's very similar, Do you think adding the communication component into that made your decision easier or harder? Q Um, it didn't, it didn't have much impact on my decision. Um. So, I don't think it made it harder. I think I was just moving the position a little bit. Um. Ye.
M4									
M5							I don't care about my communication with the doctor. Um, it limits social activities, so again, I think I will		

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			**3.1.2.1.1 Improved health state valuation**	**3.1.2.1.2 No impact on health state valuation**	**3.1.2.1.3 Worsened health state valuation**	**3.1.2.2.1 Improved health state valuation**	**3.1.2.2.2 No impact on health state valuation**	**3.1.2.2.3 Worsened health state valuation**	
							give it a 70. Its not that bad.		
M6			Oh, so this is good. So if you can have faith in your doctor and can talk through your problems then ye, that's probably, I'd probably rate it at 45, no 46, because it goes in scales of 2.					Oh, ye, communication problems with the doctors, it's really not good. I think I'll put that as 20. So, if you can't have a good consultation with the doctor, then who can you have a good consultation with?	so that's essentially, the same as the one that I've just previously done. OK, so, the consultation with the doctor, oh so I guess its going to be exactly the same again.
M7									
M8	Uhuh. Do you think having the communication added onto that make it any different? Did it make your decision any easier or harder?	Uhuh. Do you think having the communication added onto that make it any different? Did it make your decision any easier or harder? X Um, a little bit harder, because sometimes I need to. You know, this is not an easy question. You need to think a	Um, I would say 50 in the last one, because you know the doctor say um, and communication experience is good. So it					I would say 30 in this one. Um, because you know, the doctor is a professional, when after you know, after you know you went to see him and he decided to tell me that I've got some problem in me then obviously you know I may not be	

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
	X Um, a little bit harder, because sometimes I need to. You know, this is not an easy question. You need to think a little bit longer time. But it's really hard to decide anyway so. so it makes it more, more difficult with that added on.	little bit longer time. But it's really hard to decide anyway so. so it makes it more, more difficult with that added on.	may not be as bad as I think.					as good as I think.	
M9			OK, well this time you put the communication experience as good. So, you know, like I said it can change your perceptions and think					Well I guess, with the doctor, if the communication with the doctor was poor, um, it makes it, it can have more of a kind of negative effect with kind of what you perceive of your health state I guess.	

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			**3.1.2.1.1 Improved health state valuation**	**3.1.2.1.2 No impact on health state valuation**	**3.1.2.1.3 Worsened health state valuation**	**3.1.2.2.1 Improved health state valuation**	**3.1.2.2.2 No impact on health state valuation**	**3.1.2.2.3 Worsened health state valuation**	
			that, you know, because it was good then ye, your health state can be a bit better I guess.						
M10								I suppose it's the, it's the same as the last one except you've just had a rubbish experience with your doctor, and how that's made me feel, or you feel, gloomier about your, the outcome or, gloomier about how you might deal with the problems you're facing so, um. Ye, that might not be a long term kind of impact on the health state, but it can have an um, impact, on initial moods, um, so I'lll put that a little bit lower.	
M11			You have lot						

	3.1.1 Ability to score		3.1.2 Impact on score						
	3.1.1.1.Difficult to score	3.1.1.2 Easy to score	3.1.2.1 Impact of good communication			3.1.2.2 Impact of poor communication			3.1.2.3 No change to HS valuation
Transcript			3.1.2.1.1 Improved health state valuation	3.1.2.1.2 No impact on health state valuation	3.1.2.1.3 Worsened health state valuation	3.1.2.2.1 Improved health state valuation	3.1.2.2.2 No impact on health state valuation	3.1.2.2.3 Worsened health state valuation	
			of energy some of the time, its is the same but you value your communicati on as good. So I think the communicati on can up bring up my hope a little so I'll out it at 55, the same as this one, even though my physical health is worse, but probably the doctor will encourage me a bit more.						

Table 71. Framework matrix theme 4.

	4. Impact on preference for health states (TTO)	
Transcript	4.1 Process domain has no impact	4.2 Process domain has impact
F1		I'd rather have a crappy doctor for 1 week and a really bad week actually than a poor experience after 4 weeks of illness, so I'll disagree with that.
F2		
F3		
F4		
F5		
F6	no I don't see, I don't, I don't really, I wouldn't really get influenced by the type of communication I have by my doctor, and I don't think that, no I'd still feel the same things for 4 weeks, 1 week. This one for sure. But I always put like, symptoms are the way I feel. Like they are always a priority, I don't think this doctor thing has anything to do with these answers I am giving.	
F7		
F8		
F9	so when we're asking you to basically value, or say kind of gamble away time alive, do you think this is important about the consultation when you're making decisions like that? A7 No. I don't think it is no, I think you sort of draw your own experiences as well, um, it might sway you a little bit, uum, but you don't, its not really um. I don't think it swayed it at all. You know. I think sometimes they're a little bit of a red herring almost, although you do take it into account. But, um. No.	I've jumped to the last line as well and its says your communication experience was very poor, so obviously it's not looking quite as good.
M1		
M2		
M3		
M4		
M5	I don't care much for the doctor again so I won't change my decision.	
M6		
M7	it assumes that um, possibly the health state has prevented me from communicating with the doctor very well. Um, but it won't influence my decisions as much.	
M8		
M9		Ye, it's the same principle I guess, and um, this time, this time the doctor was in it and it's a was the communication was good, and B was it wasn't it was very poor, so that has an

	4. Impact on preference for health states (TTO)	
Transcript	4.1 Process domain has no impact	4.2 Process domain has impact
		added effect as well.
M10		Again, um, I feel that quality of life would be better um, for 2 years, then health state A for 5 years, especially, there's um, you'd feel more rubbish, um, if you didn't feel the doctor was supportive, um, so health state um, A for 5 years versus 1.5 years, so, I think full health for 1.5 would be better again because of quality of life, um.
M11	Then so it is the same, so the doctor doesn't really affect my decision. OK. So I think the previous one it should be this way. Because it won't affect me, it wouldn't affect me how I talk with the doctor. it wouldn't really affect me how the doctor talked to me in a sense because like you are having a longer period of time, I mean if you are going to see the doctor just for one, just for one occasion then probably he might offend you because you can't like keep in your mind, but in this case you are living in the disease for like 4 weeks, which so I suppose it wouldn't matter what the doctor says, if it were me you know. OK, so in this case. So, if the doctor treat me very poorly, would it affect my, I mean my decision to live? Um, no I wouldn't let him. I would still do what I want to do.	

Table 72. Framework matrix theme 5 .

	5.1 Impact of time spent in health state						
	5.1.1 Years (TTO)		5.1.2 Time not reported (VAS)			5.1.3 Days (TTO)	
Transcript	**5.1.1.1 Process wouldn't affect me**	**5.1.1.2 Process would affect me**	**5.1.2.1 Process wouldn't affect me**	**5.1.2.2 Process would affect me**	**5.1.2.3 Assumptions**	**5.1.3.1 Process wouldn't affect me**	**5.1.3.2 Process would affect me**
F1	I guess, I'd maybe putting it into the context of time, thinking well could I cope with a crap doctor for 5 years? But then, if you think about it, actually you cope, you'd probably change doctor if you felt that strongly about it, so maybe contextualize it in terms of time.			I guess, I'd maybe putting it into the context of time, thinking well could I cope with a crap doctor for 5 years? But then, if you think about it, actually you cope, you'd probably change doctor if you felt that strongly about it, so maybe contextualize it in terms of time. I gave a bit less thought to it because I thought, well I had to stay in that scenario. Doctor wouldn't be the main thing for me, it would be the other things that id prioritized. Whereas on a scale like that it's much more open to up or down by a certain degree, so I could say well it'll be a 70 if I had just that health state, but then with a dad doctor it would go down a bit, for a good doctor it would go up a		I guess, I'd maybe putting it into the context of time, thinking well could I cope with a crap doctor for 5 years? But then, if you think about it, actually you cope, you'd probably change doctor if you felt that strongly about it, so maybe contextualize it in terms of time. I gave a bit less thought to it because I thought, well I had to stay in that scenario. Doctor wouldn't be the main thing for me, it would be the other things that id prioritised. Whereas on a scale like that it's much more open to up or down by a certain degree, so I could say well it'll be a 70 if I had just that health state, but then with a dad doctor it would go down a bit, for a good doctor it would go up a	

	5.1 Impact of time spent in health state						
	5.1.1 Years (TTO)		5.1.2 Time not reported (VAS)			5.1.3 Days (TTO)	
Transcript	5.1.1.1 Process wouldn't affect me	5.1.1.2 Process would affect me	5.1.2.1 Process wouldn't affect me	5.1.2.2 Process would affect me	5.1.2.3 Assumptions	5.1.3.1 Process wouldn't affect me	5.1.3.2 Process would affect me
				bit, where as in terms of half years, it its almost its smaller increments than half a years' worth of compromise.		bit, where as in terms of half years, it its almost its smaller increments than half a years' worth of compromise.	
F2	Um, it didn't, it didn't have much impact on my decision. Um. So, I don't think it made it harder. I think I was just moving the position a little bit. Um. Ye.						
F3							
F4							
F5							
F6							
F7							
F8							
F9							
M1		I think it was definitely more important if it's followed by death. I think that's really important that you are having a good communication with your doctor. Um, in the other, um the other one was where you compare 2 health states and which one you would rather be					

	5.1 Impact of time spent in health state						
	5.1.1 Years (TTO)		5.1.2 Time not reported (VAS)			5.1.3 Days (TTO)	
Transcript	5.1.1.1 Process wouldn't affect me	5.1.1.2 Process would affect me	5.1.2.1 Process wouldn't affect me	5.1.2.2 Process would affect me	5.1.2.3 Assumptions	5.1.3.1 Process wouldn't affect me	5.1.3.2 Process would affect me
		in. ye, ummmmmm (silence for a few seconds, thinking) and it was only ever a month so it wasn't, again I don't think it was significantly significant for a month. Whereas the other one where you are talking about much greater expanse of time, and is followed by a definite death, I think it's really important that you have some encouragement, or that you feel like you are going somewhere with your doctor. Coz if not, that just, ye, can tip you over the edge a lot, so ye.					
M2							I think when maybe 2 weeks or 3 weeks and you've had no um, sort of positive consultation with your doctor, um, then its going to start to, um, make it a lot more different to a, um, make it a lot more serious than A,
M3			Do you think adding the communication component into that made			I suppose you might have waited a long time for your	Damn, poor or very poor on the consultation. I feel its

	5.1 Impact of time spent in health state						
	5.1.1 Years (TTO)		5.1.2 Time not reported (VAS)			5.1.3 Days (TTO)	
Transcript	5.1.1.1 Process wouldn't affect me	5.1.1.2 Process would affect me	5.1.2.1 Process wouldn't affect me	5.1.2.2 Process would affect me	5.1.2.3 Assumptions	5.1.3.1 Process wouldn't affect me	5.1.3.2 Process would affect me
			your decision easier or harder? Q Um, it didn't, it didn't have much impact on my decision. Um. So, I don't think it made it harder. I think I was just moving the position a little bit. Um. Ye.			consultation as well, and it might be only once every 6 months or so, then if its, and if it's um, a long term problem. Do you think adding the communication component into that made your decision easier or harder? Q Um, it didn't, it didn't have much impact on my decision. Um. So, I don't think it made it harder. I think I was just moving the position a little bit. Um. Ye.	slipping towards health state A being preferable for all lengths, but perhaps I've got this wrong?
M4	So, just now I don't think that would be, because its 5 years, that would be something that quickly changes, um perception. Because, um. It depends like, if that situation, in the beginning of the 5 years, and I'm um, talking to the doctor. Well my attitude, toward doctors						

	5.1 Impact of time spent in health state						
	5.1.1 Years (TTO)		5.1.2 Time not reported (VAS)			5.1.3 Days (TTO)	
Transcript	5.1.1.1 Process wouldn't affect me	5.1.1.2 Process would affect me	5.1.2.1 Process wouldn't affect me	5.1.2.2 Process would affect me	5.1.2.3 Assumptions	5.1.3.1 Process wouldn't affect me	5.1.3.2 Process would affect me
	like um, if I'm not satisfied with one, I would again try another one, so I'm not, I wouldn't be like, and I think I'm making decisions mostly by myself, so if the experience with the doctor was poor, the doctor maybe really negative, but then I'm also thinking like well, this is something about 5 years, or um, there will be many people around and, um, so, I wouldn't change much like, maybe (fills in the table), ye, the doctor wouldn't change it.						
M5							
M6		That's on the second question. Yes. Because. Ye, is the first one where you are going to recover regardless. So I think. I know 4 weeks might sound not very much for someone who hasn't personally experienced				That's on the second question. Yes. Because. Ye, is the first one where you are going to recover regardless. So I think. I know 4 weeks might sound not very much for someone who	

	5.1 Impact of time spent in health state						
	5.1.1 Years (TTO)		5.1.2 Time not reported (VAS)			5.1.3 Days (TTO)	
Transcript	5.1.1.1 Process wouldn't affect me	5.1.1.2 Process would affect me	5.1.2.1 Process wouldn't affect me	5.1.2.2 Process would affect me	5.1.2.3 Assumptions	5.1.3.1 Process wouldn't affect me	5.1.3.2 Process would affect me
		such bad health problems. It's probably going to feel like 4 years. But still its only 4 weeks and then you'll be back to normal. Whereas this one, it's very extreme, so ye, I think that in question 2, the doctor consultation would make a difference. Whereas for the first one, no.				hasn't personally experienced such bad health problems. It's probably going to feel like 4 years. But still its only 4 weeks and then you'll be back to normal. Whereas this one, it's very extreme, so ye, I think that in question 2, the doctor consultation would make a difference. Whereas for the first one, no.	
M7							
M8							
M9	I don't think the doctors like, it would kind of effect you a bit, but to be honest, um, it's hard to say what exactly happens and stuff, but if I'm in a state where I've only got 5 years to go, or I've got a quality of life for lower years I don't think the doctors kind of opinion would make a difference. Yes, because when there is death you kind of						Yes, because when there is death you kind of know. Um, its, this is kind of less, like I think the doctors kind of opinion matters more here than it does in the death kind of questions.

	5.1 Impact of time spent in health state						
	5.1.1 Years (TTO)		5.1.2 Time not reported (VAS)			5.1.3 Days (TTO)	
Transcript	5.1.1.1 Process wouldn't affect me	5.1.1.2 Process would affect me	5.1.2.1 Process wouldn't affect me	5.1.2.2 Process would affect me	5.1.2.3 Assumptions	5.1.3.1 Process wouldn't affect me	5.1.3.2 Process would affect me
	know. Um, its, this is kind of less, like I think the doctors kind of opinion matters more here than it does in the death kind of questions.						
M10					even just for a short bit, so it can have an impact, but then I suppose overall depending on the time you're struggling with something um, it might make you feel worse if your doctors kind of still, if you're still having poor experiences with them, but overall it might not have too much impact in the end because um, you still dealing with all of these problems and your doctor's kind of one part of them and I really feel that kind of lone, I can't think of the word now, hopelessness and stuff, when you first see your doctor, but then I've kind of put that out of		

	5.1 Impact of time spent in health state						
	5.1.1 Years (TTO)		5.1.2 Time not reported (VAS)			5.1.3 Days (TTO)	
Transcript	5.1.1.1 Process wouldn't affect me	5.1.1.2 Process would affect me	5.1.2.1 Process wouldn't affect me	5.1.2.2 Process would affect me	5.1.2.3 Assumptions	5.1.3.1 Process wouldn't affect me	5.1.3.2 Process would affect me
					my head after a while, because it seems to me, well you might seem them more frequently but for me it's been less regular. (assumed short term) Ye, that might not be a long term kind of impact on the health state, but it can have an um, impact, on initial moods, um, so I'lll put that a little bit lower. (assumed short term)		
M11					Because my disease is only a one off, so if you want to treat me well then, that's fine. If he doesn't treat me well because of that. I mean if he doesn't communicate me well just because I feel bad and then just that disease takes to be longer, so I was thinking, I was thinking with the communication effect	it wouldn't really affect me how the doctor talked to me in a sense because like you are having a longer period of time, I mean if you are going to see the doctor just for one, just for one occasion then probably he might offend you because you can't like keep in your mind, but in this case you are living in the disease for like 4	

	5.1 Impact of time spent in health state						
	5.1.1 Years (TTO)		5.1.2 Time not reported (VAS)			5.1.3 Days (TTO)	
Transcript	5.1.1.1 Process wouldn't affect me	5.1.1.2 Process would affect me	5.1.2.1 Process wouldn't affect me	5.1.2.2 Process would affect me	5.1.2.3 Assumptions	5.1.3.1 Process wouldn't affect me	5.1.3.2 Process would affect me
					your disease recovery time, but in this case, since you define a time, then it doesn't really matter much. (assumed one occasion) I think so, but in that case, because of the one occasion, it would affect me, I mean personally a bit more than this. (assumed one occasion)	weeks, which so I suppose it wouldn't matter what the doctor says, if it were me you know. OK, so in this case. Because my disease is only a one off, so if you want to treat me well then, that's fine. If he doesn't treat me well because of that. I mean if he doesn't communicate me well just because I feel bad and then just that disease takes to be longer, so I was thinking, I was thinking with the communication effect your disease recovery time, but in this case, since you define a time, then it doesn't really matter much. Right, and what about the 2 types of question here? So one was asking you about 2 health states, and one was asking you about	

	5.1 Impact of time spent in health state						
	5.1.1 Years (TTO)		5.1.2 Time not reported (VAS)			5.1.3 Days (TTO)	
Transcript	5.1.1.1 Process wouldn't affect me	5.1.1.2 Process would affect me	5.1.2.1 Process wouldn't affect me	5.1.2.2 Process would affect me	5.1.2.3 Assumptions	5.1.3.1 Process wouldn't affect me	5.1.3.2 Process would affect me
						time alive basically, do you think communication had um the same impact in both of those or different? Ye, to me it doesn't really matter, so in this case my answer is still the same, in a sense, so to me it doesn't matter to me in this particular case.	

Table 73. Framework matrix theme 6 (Section 6.1).

6.1 Relative Importance of process domain				
Transcript	**6.1.1 Least important**	**6.1.2 Second importance**	**6.1.3 Middle importance**	**6.1.4 Important**
F1	Um, and the doctor, I think adds a huge influence on all of those, on, especially on the kind of mentality of a person when they go into a doctor if you're feeling low and your doctor picks you up that has an impact on how you approach your disease or. So priority wise I'd say activity and social, um, social activities and mental state are the most important for me, rather than work and housework and things like that, and people at work will always help you get through physical or emotional problems, you know. There's things in place for that but really I don't know, I class those as being more important than going to the doctor. Doctor wouldn't be the main thing for me, it would be the other things that id prioritised.			
F2		And then the hope of the doctor.		
F3				
F4		Um, ok. So I think the downhearted the most. And then, um, the communication if it is an issue. The least I think the work and social activities. Ye.		
F5				
F6	This one for sure. But I always put like, symptoms are the way I feel. Like they are always a priority, I don't think this doctor thing has anything to do with these answers I am giving.			
F7				
F8	The pain was most important. Because it would really affect your quality of life. Um, and having no energy			

6.1 Relative Importance of process domain				
Transcript	**6.1.1 Least important**	**6.1.2 Second importance**	**6.1.3 Middle importance**	**6.1.4 Important**
	and having no, no being tense and downhearted. The consultation with the doctor, it didn't affect the quality of life, although it probably would have made you feel a bit low. I don't think it would have affected it much. It was more the pain that I was interested in, and the activities on some of them it said you couldn't go out at all, you couldn't do anything. Whereas the communication with the doctor was poor I don't think, it would have mattered so much, because that's just a small part of your life, the consultation with your doctor. This bit of information about the consultation, do you think it's a good idea, or not a good idea? A6 Um, I don't think it's, when you're looking at the wider context of your health. I know it's disappointing to have a poor consultation with your doctor. But I think in the wider scheme of things I'm only interested in the actual health. Um. I mean it's a plus if you have a good consultation but I don't think that it can't all be good and if you've got a poor health outcome anyway. If you have a good consultation with the doctor it's not going to do anything to change your health. You can't do anything about your health anyway. So, I don't think it's, it's that important. I think it's all of the other factors that were important.			
F9	I'm not really taking that into consideration really. I'm not really looking at that one. No, no, I'm just sort of taking the, the other questions into account, not the one with the doctor.			
M1	Definitely the tension, and feeling downhearted all of the time. That sort of thing. So to me, the emotional state is more important than, it kind of outweighs any physical pain that you might be feeling moderately or quite a bit. So those were more important. And if it interferes with your social activities as well, the pain			

6.1 Relative Importance of process domain				
Transcript	**6.1.1 Least important**	**6.1.2 Second importance**	**6.1.3 Middle importance**	**6.1.4 Important**
	that you are feeling. I think those two were most important. For me, it wasn't particularly significant because, ye, you can go to another doctor. So it's not so bad. And especially since there is a lot of other help available, like you can use the internet, you can talk to people, you can find out different ways, for you dealing with it. So it's not. For it wasn't particularly significant. But I guess, it is more significant if you are feeling tense already and then you go to the doctor and don't get anything from that, and it is a very poor consultation. Then it adds there. But if you are feeling okish and you have a poor consultation. I still think you would get over it. Do you know what I mean?			
M2				And so, when you were looking at the health states, which aspects of it were most important to you when you were coming to your decisions? P OK, always the doctor. it's extremely important because when you are suffering from all of those things people look to an expert to tell them, um, what to do, or what they can do to make all of those things better. So ye, I um, I think the word I used was hope wasn't it. So there you go. I think it is very very important, probably more so than the rest of the things.
M3				
M4				
M5				
M6			Which ones? The pain one definitely. I would definitely go for pain. Um, energy.	I would rate it fairly high.

6.1 Relative Importance of process domain				
Transcript	**6.1.1 Least important**	**6.1.2 Second importance**	**6.1.3 Middle importance**	**6.1.4 Important**
			That's number 2 maybe. Lets see. Maybe, number, oh, no problems with work or regular daily activities as my third one. Ye, that's the order to go in, and the rest of them are not too concernded about. VB Where in that order would the consultation with the doctor come? Lets see. I'd probably put it as number 4. Ye, so just after the 3 I've just mentioned.	
M7				
M8		Um, and where abouts within that would you place the communication one in the importance of things? X (Silence for 6 seconds) (Sigh) I would say second. So, for number 4 (pain) third) and second, and 5 on the top. VB OK, so 5, 7, 4 yep. X Because like number 4 (pain) if you are feeling depressed like you don't even want to see the doctor. And then, I'd put number 7 the second one, like if you are ok then you can still see the doctor, and the doctor tell you what happen, and you can get any treatment. And then I would put 4 third, because it doesn't stop me from going to see the doctor. So I feel pain but I can still go		

6.1 Relative Importance of process domain				
Transcript	**6.1.1 Least important**	**6.1.2 Second importance**	**6.1.3 Middle importance**	**6.1.4 Important**
		to see the doctor. But if I'm feeling depressed um, I don't want to do anything. I just want to sit at home and cry, and don't want to do anything.		
M9	And ye, I'd say the consultation with the doctor is last. Hm. Um. Again it's like, you don't know exactly what happens, but it, you know its still kind of a factor that will affect like your, like, kind of you know, what you perceive as your health state I guess. It is kind of important, but um. Umm, its, I wouldn't give it you know, the highest how I would priorities it.			
M10			I'd say um, feeling tense or downhearted and having energy or not maybe, um, because they kind of impact on all of the others. Um, and pain. Then I'd maybe say the doctor, um, and then the other ones.	
M11	it's generally my health state rather than a communication with the doctor. So communication with the doctor is definitely not the most important one, because I can't feel like my attitude to live towards life or towards my health is the most important. So basically, if I'm generally healthy its much better. VB OK, so the communication you saw as least important? A3 Ye, I would put it last.			

Table 74. Framework matrix theme 6 (Section 6.2).

	6.2 Importance of process domain (not relative)					
Transcript	**6.2.1 When health states similar**	**6.2.2 Not-important**	**6.2.3 Increased with questions**	**6.2.4 Important**	**6.2.5 Dependent on which valuation method**	**6.2.6 Dependent on severity of health state**
F1						
F2						
F3			I think at first, I kind of said to you, I kind of dismissed it a bit I think, and said well you can speak to your partner and so what if you have a bad kind of consultation, um, but I think, the more I read it, the more I thought ye, that would really impact on how I felt. Um, ye, I could. OK, right, and do you think that 7th information about communication, do you think it was more or less important in any of the types of questions? V Ye, as it went on, um, I do. I think that became that influenced definitely how I was thinking about the health states more and more.			
F4				VB And was the information about the communication with the doctor as important to you in both? Y Um, yes, I think so. I think it depends on the patient.		

	6.2 Importance of process domain (not relative)					
Transcript	**6.2.1 When health states similar**	**6.2.2 Not-important**	**6.2.3 Increased with questions**	**6.2.4 Important**	**6.2.5 Dependent on which valuation method**	**6.2.6 Dependent on severity of health state**
F5				Um, you've just had a con… ye, I think that's quite, quite significant, that you've tried to explain to your doctor how you're feeling and um, what, I'm assuming, or maybe I'm assuming that, um, you haven't communicated very well, but it could be that the doctor possibly hasn't received what you are saying very well. Because I would think not being able to explain to the doctor how you feeling is quite significant really. I mean I think that it is major, because I know myself, if do go to the doctor, and I feel I've been able to express what I want to say well, and um, I have a fantastic GP, and I just always feel better (giggles) when I've seen him. Um, but I, um. (Slaps lips). I never have been in the situation where I've felt I could express. Um, well I think I've just said something about that haven't I. That um, being able to have good communication both ways with the doctor, um, does have an um, a profound effect on how you feel about um, your health state.		

	6.2 Importance of process domain (not relative)					
Transcript	**6.2.1 When health states similar**	**6.2.2 Not-important**	**6.2.3 Increased with questions**	**6.2.4 Important**	**6.2.5 Dependent on which valuation method**	**6.2.6 Dependent on severity of health state**
F6		VB OK, and you obviously didn't think this was important at all. The communication. Why? A4 Because, I don't know. I maybe its just the kind of person I am, or just, um. I much more rely on what I feel like and on my personal evaluation of myself. Even though that might sound a bit, I don't know. Um. Too confident or assertive. But still I, I don't think that the type of communication with my doctor would be something which would determine anything in my health state. I don't, I just don't find it important.				
F7	some of them it didn't make a difference, but a few of them, where they were quite similar, but it was, additionally they said they had communicated very poor, it stood out that, ok you need to take that into consideration.			I think the last thing that's key here is the last bullet point, about communicating with your doctor		
F8		This bit of information about the consultation, do you think it's a good idea, or not a good idea? A6 Um, I don't think it's, when you're looking at the wider context of your health. I know it's disappointing to have a poor				

	6.2 Importance of process domain (not relative)					
Transcript	**6.2.1 When health states similar**	**6.2.2 Not-important**	**6.2.3 Increased with questions**	**6.2.4 Important**	**6.2.5 Dependent on which valuation method**	**6.2.6 Dependent on severity of health state**
		consultation with your doctor. But I think in the wider scheme of things I'm only interested in the actual health. Um. I mean it's a plus if you have a good consultation but I don't think that it can't all be good and if you've got a poor health outcome anyway. If you have a good consultation with the doctor it's not going to do anything to change your health. You can't do anything about your health anyway. So, I don't think it's, it's that important. I think it's all of the other factors that were important.				
F9						
M1		For me, it wasn't particularly significant because, ye, you can go to another doctor. So it's not so bad. And especially since there is a lot of other help available, like you can use the internet, you can talk to people, you can find out different ways, for you dealing with it. So it's not. For it its wasn't particularly significant. But I guess, it is more significant if you are feeling tense already and then you go to the doctor and don't get anything from that, and it is a very poor consultation. Then it adds there. But if you are feeling okish and you have a poor consultation. I still think you would get over it. Do you know what I mean?		I think it was definitely more important if it's followed by death. I think that's really important that you are having a good communication with your doctor. Whereas the other one where you are talking about much greater expanse of time, and is followed by a definite death, I think it's really important that you have some encouragement, or that you feel like you are going somewhere with your doctor. Coz if not, that just, ye, can tip you over the edge a lot, so ye.		

	6.2 Importance of process domain (not relative)					
Transcript	**6.2.1 When health states similar**	**6.2.2 Not-important**	**6.2.3 Increased with questions**	**6.2.4 Important**	**6.2.5 Dependent on which valuation method**	**6.2.6 Dependent on severity of health state**
M2				I think I mentioned sort of previously what I think the consultation with the doctor is an important thing, Again, I think um, the stuff about having a consultation with a doctor, the expert, on what you are suffering, is the key, and I think um, although the um, issues are lesser with A than they are with B, they are still severe, so having a negative experience with the doctor, or a non-productive experience with the doctor is um going to heighten that more for you, its not going to make you feel better about it than, (slaps lips), so, I was going to, I think they are both essentially the same as each other. And so, when you were looking at the health states, which aspects of it were most important to you when you were coming to your decisions? P OK, always the doctor. it's extremely important because when you are suffering from all of those things people look to an expert to tell them, um, what to do, or what they can do to make all of those things better. So ye, I		

	6.2 Importance of process domain (not relative)					
Transcript	**6.2.1 When health states similar**	**6.2.2 Not-important**	**6.2.3 Increased with questions**	**6.2.4 Important**	**6.2.5 Dependent on which valuation method**	**6.2.6 Dependent on severity of health state**
				um, I think the word I used was hope wasn't it. So there you go. I think it is very very important, probably more so than the rest of the things.		
M3		And again, I don't know how much value I value on the consultation evaluation. I mean, I don't think I'd be relying on that. I don't think I'd be. I don't think I'd be looking. I mean I guess some people would be looking forward to having a good outcome from that. Sorry a good evaluation, coming to a good conclusion.				
M4						
M5		I don't care about my communication with the doctor. Um, it limits social activities, so again, I think I will give it a 70. Its not that bad. And I don't care about communication with the doctor. I' still don't care about the doctor I don't care much for the doctor again so I won't change my decision. Um, the most important part was um, the social activities, um, and the work things, so if it hampers my work and if it's due to emotional problems then that's even worse, and if it hampers my				

	6.2 Importance of process domain (not relative)					
Transcript	**6.2.1 When health states similar**	**6.2.2 Not-important**	**6.2.3 Increased with questions**	**6.2.4 Important**	**6.2.5 Dependent on which valuation method**	**6.2.6 Dependent on severity of health state**
		social activities then that's even more, because you know, hm, I don't care about the housework that much. I don't care about if I feel downhearted some times, or um, it affects me more if it's like physical things, and I don't care anything about the doctor, or the energy. it's not a major part of anyone's life, so, it's not, it's not, it's only a fraction of the experience that you have. The doctor? No, basically I care as much as if a sales cleric has been nice to me. Why would I be bothered? It doesn't seem that important to me, it just seems like a small section of, it's like for me a doctor is just a person who services your health. You go, you use his services, you should be happy from the way he works, but it shouldn't, he's not a councillor, he's not, he's not obliged, it's a good thing if he helps you emotionally, if he gives you help, but it's not what he is paid to do.				
M6						So the difference between having a doctor and not having a doctor might be minimal, or it could be a big difference. It really depends. Its really hard

	6.2 Importance of process domain (not relative)					
Transcript	**6.2.1 When health states similar**	**6.2.2 Not-important**	**6.2.3 Increased with questions**	**6.2.4 Important**	**6.2.5 Dependent on which valuation method**	**6.2.6 Dependent on severity of health state**
						to say. I'd probably say it might help a little.
M7						
M8				Definitely it is important communication, like because the doctor is a professional, like if they give you like some advice then you, like it may change their mind, or you may treat, like you may feel like the disease or you know more serious. Like sometimes you never know how suffering you are until your doctor tell you Uh, uh. Do you think this is as important the communication, for each of these?		
M9						
M10		I don't think it had too much of an impact, um, ye because the overall conditions I suppose were what I compared more than the bit with the doctor.			well in terms of comparing 2 health states I think it made a difference. But when you're comparing to full health, I don't think it did make such a difference.	
M11						

Table 75. Framework matrix theme 7.

	7.1 Positive view			7.2 Negative view	7.3 Relevant for some	7.4 Distinction between process and health made
Transcript	**7.1.1 Important information to include**	**7.1.2 Makes sense**	**7.1.3 Other**			
F1			OK so I guess it's useful in that way to kind of set on a scale how that would affect the persons perceived health state, then I think that's a good idea. It's just, how it's used I guess is the, the thing. But ye, doctors like individuals doctors are individuals and they might have had a bad day, and just made that experience poor for that patient. So, I guess that's maybe something to, account for, if you're asking what I feel about that statement, I think its subject to individuals and that's the problem with that method of maybe measuring, um although I can say if I had a good example of a doctor and a bad example of a doctor I would feel differently about the health state I was in at the time.	But ye, doctors like individuals doctors are individuals and they might have had a bad day, and just made that experience poor for that patient. So, I guess that's maybe something to, account for		
F2	I think it gives people a different view, definitely. Because I think, if that's not there. Most people would just assume you would have a positive experience, or you					

	7.1 Positive view			7.2 Negative view	7.3 Relevant for some	7.4 Distinction between process and health made
Transcript	**7.1.1 Important information to include**	**7.1.2 Makes sense**	**7.1.3 Other**			
	could go and see someone, or something could be done, you know. But I think if that's put in place, and it's, you know. They say, oh um, you had a bad experience or a bad experience, that sort of changes, it would change my view definitely. I think a lot of people do have bad experience with GPs and with the NHS in general, but without that there, you need to know there is someone you can go to and that you know, well you need to because where would you go otherwise? Like you said, ok just the process of, even putting the GP into the mix, and just giving them the option. Because if you put you feel this, you feel low, you feel this, and they'd think oh ye I do. But if you put you feel this, you feel this, but then GP you'd be like oh, I can go see my GP about that, even you know just that like link, even if it wouldn't mean that, just putting that in there gives them something else to think about, and to give them that other option.					

	7.1 Positive view			7.2 Negative view	7.3 Relevant for some	7.4 Distinction between process and health made
Transcript	**7.1.1 Important information to include**	**7.1.2 Makes sense**	**7.1.3 Other**			
	Because depressed people sometimes don't even see what's around them.					
F3						
F4					VB So, do you think it makes sense including this, or not? Y Um, I think it's a case of, I think it depends on the patient. If the doctor feels like the patient is open and honest because I think doctors can read patients, they deal with, they work with people and deal with people, so they can a lot of the time tell who is hiding things and who is not. So I think it would depend on a case by case basis, you know where I think where I'm quite an open person, and I think that comes across, so I think a doctor saying that to me would probably be more worthwhile than saying it to my partner who would you know, go away, you don't know what you're talking about, I'm right, I'm always right, who are you?	

	7.1 Positive view			7.2 Negative view	7.3 Relevant for some	7.4 Distinction between process and health made
Transcript	**7.1.1 Important information to include**	**7.1.2 Makes sense**	**7.1.3 Other**			
					So you know. I think it would depend on the patient.	
F5						
F6						you value your communication experience as good. Still, but I still have the same symptoms though? Ye, same symptoms, no. That's. OK. I don't think that the type of communication with my doctor would be something which would determine anything in my health state.
F7						
F8						
F9			Do you think it altered your value? The way you valued that health state? A7 Ye, definitely, um, because, I'm I think as I was going along I was sort of jumping to the top and then jumping to the bottom to see what the outcome was, a little bit more for the doctor. And thinking oh alright, this isn't going to be as well, or this is going to be better this one. So it sort of surmised it			

	7.1 Positive view			7.2 Negative view	7.3 Relevant for some	7.4 Distinction between process and health made
Transcript	**7.1.1 Important information to include**	**7.1.2 Makes sense**	**7.1.3 Other**			
			slightly.			
M1						
M2			It's extremely important because when you are suffering from all of those things people look to an expert to tell them, um, what to do, or what they can do to make all of those things better. So ye, I um, I think the word I used was hope wasn't it. So there you go. I think it is very, very important, probably more so than the rest of the things.			
M3		Oh, ok. Um I don't know, I think um, I think they were easy to understand. Once you told me that they were.				
M4	Well, when it's good, it says good, and its good, I would like to reflect that somehow.			Well, when it's good, it says good, and its good, I would like to reflect that somehow. Um, because I have problems like, the case it wasn't good. Um, I think the worst one is like, was it helpful or not, but the other way if you ask it, was it very poor? People might not, I mean I wouldn't go, I would just go to another one, another doctor. I wouldn't bother.	Um, I think the worst one is like, was it helpful or not, but the other way if you ask it, was it very poor? People might not, I mean I wouldn't go, I would just go to another one, another doctor. I wouldn't bother.	

	7.1 Positive view			7.2 Negative view	7.3 Relevant for some	7.4 Distinction between process and health made
Transcript	**7.1.1 Important information to include**	**7.1.2 Makes sense**	**7.1.3 Other**			
M5				VB OK, um, and, as far as the doctor that you say is not important to you at all, so do you think that it's important to include it in this at all or not. T No, I don't think so, because it's only a, how often does a person really like to except when he is in a hospital, you know like it's not a major part of anyone's life, so, it's not, it's not, it's only a fraction of the experience that you have. The doctor? No, basically I care as much as if a sales cleric has been nice to me. Why would I be bothered?		
M6					Um, not really, because its not very clear what the problem is, I mean I don't know based on its context how treatable their problem is? So the difference between having a doctor and not having a doctor might be minimal, or it could be a big difference. It really depends. It's really hard to say. I'd probably say it	

	7.1 Positive view			7.2 Negative view	7.3 Relevant for some	7.4 Distinction between process and health made
Transcript	**7.1.1 Important information to include**	**7.1.2 Makes sense**	**7.1.3 Other**			
					might help a little.	
M7						
M8						
M9						
M10				I don't think it had too much of an impact, um, ye because the overall conditions I suppose were what I compared more than the bit with the doctor.		
M11					Um, I think to some people it does make a big difference. For, I mean, I mean for me, a person who still has his own family and all, so um, I've not married or anything. I mean this demographic, to me however the doctor talk to me it doesn't matter, because if he doesn't want to treat me more, that's fine. (Giggles) So long as he does his work and then doesn't kill me its fine. I mean if you take like a man who is 80 years old, generally, or a lady who is quite fragile frail then the doctor in front of them don't care about what they're saying "I'm in pain" "I don't care" I mean your disease is more important.	

	7.1 Positive view			7.2 Negative view	7.3 Relevant for some	7.4 Distinction between process and health made
Transcript	**7.1.1 Important information to include**	**7.1.2 Makes sense**	**7.1.3 Other**			
					So, If that the case, then um, I will feel a bit more empathetic towards them. Like doctors communication towards I mean for them its, it is important so it is really based on the age the social circumstances and to some people I would say it really it make a big difference.	

Table 76. Framework matrix theme 8.

Transcript	8.1 Heuristics
F1	
F2	
F3	
F4	
F5	
F6	
F7	
F8	
F9	I've jumped to the last line as well and its says your communication experience was very poor, so obviously it's not looking quite as good. I'm I think as I was going along I was sort of jumping to the top and then jumping to the bottom to see what the outcome was, a little bit more for the doctor. And thinking oh alright, this isn't going to be as well, or this is going to be better this one. So it sort of surmised it slightly
M1	
M2	
M3	Because my disease is only a one off, so if you want to treat me well then, that's fine. If he doesn't treat me well because of that. I mean if he doesn't communicate me well just because I feel bad and then just that disease takes to be longer, so I was thinking, I was thinking with the communication effect your disease recovery time, but in this case, since you define a time, then it doesn't really matter much. (assumed one occasion) I think so, but in that case, because of the one occasion, it would affect me, I mean personally a bit more than this. (assumed one occasion)
M4	
M5	
M6	
M7	
M8	
M9	
M10	
M11	

Table 77. Framework matrix theme 9. (Section 9.1)

	9.1.1 Positive views on technique	9.1.2 Negative views on technique			9.1.3 Interviewer prompts
Transcript		**9.1.2.1 Own concerns when completing**	**9.1.2.1 General views on technique**	**9.1.2.2 Other**	
F1	OK, I think I just read aloud, which is quite annoying, um, but I found it ok to kind of go through what I was thinking about things,	OK, I think I just read aloud, which is quite annoying, um, but I found it ok to kind of go through what I was thinking about things,			
F2	So, how generally did you find the think aloud exercise? S Ummm, uh, ye, it is hard because you sometimes think, you know what's going to come out of my mouth, but um, ye, I think when you get used to it (giggles)	So, how generally did you find the think aloud exercise? S Ummm, uh, ye, it is hard because you sometimes think, you know what's going to come out of my mouth, but um, ye, I think when you get used to it (giggles) Ye, because, because, I was reading, I was conscious that I didn't want to say out loud what I was reading, because you already know it, do you know what I'm mean, I didn't want to verbatim, but then I was rushing through because I didn't want to waste your time. No, but do you know what I mean? So I was just a bit, so ye, the questions weren't hard it was just because I couldn't switch between (giggles) reading and saying.			

	9.1.1 Positive views on technique	9.1.2 Negative views on technique			9.1.3 Interviewer prompts
Transcript		**9.1.2.1 Own concerns when completing**	**9.1.2.1 General views on technique**	**9.1.2.2 Other**	
F3	Um. I quite enjoyed it actually, ye, it was quite good trying to think um, trying to explain what I was thinking.	kind of explaining what I was thinking made me maybe a bit more cautious and take a bit more care over what I was choosing as my answer.	Sometimes its difficult to explain certain things, um, maybe if they were quite personal of if you're thinking about a certain, um, personal sort of story in your life. Then sometimes things were quite difficult to express.		
F4					
F5	what about if I hadn't been here listening to you thinking out loud. Do you think things would have been any different or the same? A1 No, I don't think so, it's actually quite nice to have somebody sitting next to you, and no, (giggles)				Ye, sorry, I was trying to think out.
F6	Oh, I think it was good. But I didn't really see it. I mean the same would have happened if I'd been thinking in my head.			Yes, except I wouldn't have seen the differences, before. But the answers would have been the same.	
F7		It's kind of strange at first. Kind of bombarding you with my thoughts, and should I keep them to myself?	It's kind of strange at first.		
F8		When you're thinking aloud it's quite difficult to think which way around. I knew what I meant in my mind. But seeing it on paper and getting it the right way around it was one of the most difficult ones.	I think it does detract, um as it's just shown, I don't think you read the question properly. Um, it's like in an exam where you read the question properly, and I think it detracts when you're trying to think aloud. I think you		

	9.1.1 Positive views on technique	9.1.2 Negative views on technique			9.1.3 Interviewer prompts
Transcript		**9.1.2.1 Own concerns when completing**	**9.1.2.1 General views on technique**	**9.1.2.2 Other**	
			don't digest the question. Especially as they're quite complicated because you've got to think of the different scenarios and you've got to think which way around they are, I think it is distracting. I think you may answer differently because you're not fully concentrating. I think it's difficult to read and digest, and talk at the same time.		
F9		I don't know, its um just weird isn't it, and I know the police do this don't they when they're driving and things. Ye, they're trained for it, you're not trained for that are you so it doesn't come naturally. You know you're trained to keep your thoughts in, you know, so just, you know. Because you think oh gosh, I might say something silly, you know. Which I may have anyway, so (giggles).	Its hard, ye it is hard, um ye, I don't know, its um just weird isn't it, and I know the police do this don't they when they're driving and things. Ye, they're trained for it, you're not trained for that are you so it doesn't come naturally. You know you're trained to keep your thoughts in, you know, so just, you know. Because you think oh gosh, I might say something silly, you know. Which I may have anyway, so (giggles).		
M1	Um, not that bad actually. I do it probably more often than I should.				
M2				Um, quite difficult, um, just personally I just find it normally when I think about things I am silent and I am thinking about things for a long time, so it is completely the opposite to how I normally think, so um, I found that kind	

	9.1.1 Positive views on technique	9.1.2 Negative views on technique			9.1.3 Interviewer prompts
Transcript		**9.1.2.1 Own concerns when completing**	**9.1.2.1 General views on technique**	**9.1.2.2 Other**	
				of made me do something which I was quite uncomfortable with, which is almost trying to quantify all of those different um, ailments, which you associate with all sorts of things like physical limitations, depression, all sorts of things, I found that all sort of, um difficult to do in a short space of time.	
M3	Uhh, (sighs), kind of, if I was to assess myself I'd say I didn't think out loud enough, I think. Well I was my struggle with it was just getting my head around the, the way that the question worked, and so I was, I don't know, kind of, I don't know going around in circles a little bit trying to work it out. But um, I don't mind, I mean I didn't mind the extent to which I thought aloud I thought that was fine.			No, I feel I might be prejudiced if I'm honest. My deepest thoughts might come out I suppose. We'll see how it goes. Uhh, (sighs), kind of, if I was to assess myself I'd say I didn't think out loud enough, I think. Well I was my struggle with it was just getting my head around the, the way that the question worked, and so I was, I don't know, kind of, I don't know going around in circles a little bit trying to work it out. But um, I don't mind, I mean I didn't mind the extent to which I thought aloud I thought that was fine.	
M4	Um, it was well, maybe because I've done consultancies it was easy. I could say everything that I thought.				
M5	Um, it was good. Yep. VB	Um, it was good. Yep. VB			

	9.1.1 Positive views on technique	9.1.2 Negative views on technique			9.1.3 Interviewer prompts
Transcript		**9.1.2.1 Own concerns when completing**	**9.1.2.1 General views on technique**	**9.1.2.2 Other**	
	Yep. T Um, I don't think it helped me that much in the decision making.	Yep. T Um, I don't think it helped me that much in the decision making.			
M6		You have to think really hard, and when I have to talk when I don't really want to it messes with me a little bit. I find it easier to work thinking silently really.	You have to think really hard, and when I have to talk when I don't really want to it messes with me a little bit. I find it easier to work thinking silently really. Really weird. It doesn't feel right. I don't know. I mean. It feels like I still have thoughts inside my head that I'm not saying out loud. It feels as though my voice is clashing with the thoughts inside my head. And I mean t hats the best way to describe it. Even though I'm saying things out loud, I'm only doing it for the purpose of this experiment. So sometimes I feel like I'm having to speak for the sake of speaking. So ye, it is a strange experience.		
M7	I think possibly it may have, it may have helped me um, think about things, a bit more laterally and a bit more holistically really.	If I maybe I didn't think aloud I may have just answered it a bit more quickly.	It was, strange initially		
M8	I think its ok. I really enjoy it				
M9	Um, it was alright, I've not				

	9.1.1 Positive views on technique	9.1.2 Negative views on technique			9.1.3 Interviewer prompts
Transcript		**9.1.2.1 Own concerns when completing**	**9.1.2.1 General views on technique**	**9.1.2.2 Other**	
	really done it before I guess, but it was normal.				
M10	it was quite interesting trying to like trying to figure out what thought process are happening, whether you were how much you were saying them all out loud.	I felt kind of conscious that a thought might have just occurred and I might not have said it sort of thing. So, ye, um, so ye that was going through my head while I was thinking certain things. Um, I suppose because it is a bit of a kind of, you know I was trying to focus on what I was thinking and saying it out loud, rather than just having them, so it, so that kind of interferes with it, so you knew what I was thinking but, I don't know, I'm just looking at what I've put. Ye. I don't know, it probably would have been quite similar.			
M11	Its good. If only I was allowed to do it in an exam. Ye, and I think ye, because when I speak out loud. I mean I cannot listen to it, if I think silently I mean really quiet I get distracted by things around me, like if there is a clock ticking. but I prefer saying it out loud.				

Table 78. Framework matrix theme 9 (Sections 9.2 and 9.3).

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
F1		Ye, without any doubt, I think the only thing I kind of avoided doing was swearing, but everything else was just what was in my mind you know.		I would probably have sea sawed between the questions a bit more if you hadn't helped me with the clarification of some bits. But that wouldn't have affected the output of what id said particularly. Like on the scale. I didn't get the concept this was a whole health state. But obviously, once I knew that, then that made it much clearer.				
F2					Um. I think they might have been better constructed answers (giggles), um, because I do think if you're asked to think out loud, then um, people know your thought processes and how you arrive at an um		Um, yes, I think so, but also having you there, um even thought I wasn't bouncing my heads off you, because you weren't obviously speaking back in that way. But obviously having a	

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
					conclusion, and you know, sometimes that's not always the least embarrassing way. But you know, my answers would have been more constructed. But they would probably have been the same.		face to sort of recognise and you know, like talk to, um, gives you that bit of flow, and helps you speak more easily. Definitely.	
F3								And what about if I hadn't been sat here listening, do you think it would have altered how you'd answered? V If I'd still had to think outloud but you weren't here? VB Ye V So if I was in a room and just kind of talking? I think that would have been different, I think it was helpful that you were here. I felt kind of prompted and I felt um I could, um, I

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
								felt like I needed to, really express myself, where as if you weren't here the less I had to talk out loud. I think I would have been quieter.
F4					So do you think if you hadn't been asked to think aloud, do you think your answers would have been the same or different? Y Um, maybe would have spent some more time thinking. Hm, maybe having to verbalise everything made me rush through it.			VB And do you think, so there is one thing thinking outloud, but there is also me sitting her listening. Do you think if I hadn't been here your answered would have been any different? Y Um, I think, well I think always having somebody in the room, you are all. I am very much aware of my surroundings all of the time. So even though I feel comfortable and relaxed. It couldn't help but influence it,

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
								maybe just a bit. I mean just being more conscious about what I'm saying and how I'm saying it, you know. Rather than the gibberish that I usually come out with. Actually forming full sentences and you know, and not talking bollocks. So um, so I think, but then I do suffer from a bit of self-consciousness yousee and my hands are a little clammy, and I am, so I do have like a nervous streak in me so I think ye. Rather than totally alone. I think it would have affected me, maybe just a little.
F5	do you think if I hadn't asked you to think aloud, do you	But do you think, um, do you think your, so your					I think they would have been the same.	

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
	think your answers would have been the same? A1 Um, I, cant, another one that's quite hard for me to answer, because I think, um part of the thinking, part of the thinking process, is there's a sort of little bit of panic. And you can't your can't quite, so you don't express that, except for once or twice, when I've said am I don't this is right. Um, I don't think um, oh.	personal answers would have been any different, or would they have been the same. A1 I think they would have been the same.					what about if I hadn't been here listening to you thinking outloud. Do you think things would have been any different or the same? A1 No, I don't think so, it's actually quite nice to have somebody sitting next to you, and no, (giggles)	
F6		Oh, I think it was good. But I didn't really see it. I mean the same would have happened if I'd been thinking in my head. Yes, except I wouldn't have seen the differences,		Yes, except I wouldn't have seen the differences, before. But the answers would have been the same.			VB Yep, ok. And what about if I hadn't been sat here listening do you think your answers would have been the same? A4 Yes.	

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
		before. But the answers would have been the same.						
F7			I think my answers are more thoughtful and more, ye accurate.		Um, I don't think they would have been as thoughtful, um. Because I might have felt like you're rushing or waiting for me to like give and answer		It might have felt a time pressure if you weren't here, like she might be back here, and it was good to have assurance. Like that I was going at it the right way. So I felt like my answers were true because of that.	
F8	I think you may answer differently because you're not fully concentrating.						I think they would have been the same, because I, I was aware if was being taped, and if it hadn't been taped I might have perhaps had longer gaps and I might have tried to digest the question more. But I was conscious that I didn't want to have big gaps. But you being sat there I don't think. But it	

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
							would good to have the prompts.	
F9		Um, possibly not, um, or obviously the ones where I'd gone wrong I wouldn't have had you there to prompt me so it probably would have pointed me to a different picture if I'd not been prompted there, but I think in general. Saying generally, no, I'd still probably have come to the same conclusions.		Um, possibly not, um, or obviously the ones where I'd gone wrong I wouldn't have had you there to prompt me so it probably would have pointed me to a different picture if I'd not been prompted there, but I think in general. Saying generally, no, I'd still probably have come to the same conclusions.			I think I would have come to the same conclusions. I think I've got quite definite ideas on, on the matter.	
M1		I think so, I think so ye.					Um and what about if I hadn't been sat in the room listening, do you think your answers would have been the same O Yes.	
M2		I think, yeee. Um, yes I do Um, ye, what id say					The same	

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
		is that everything I said would be the same, thinking out loud or not thinking out loud, on the assumption that the person I am considering is somebody who sees the doctor as the only source of information or the best source of information.						
M3		Um, the same, is that alright? No, I mean I asked a couple of questions of you, and so I think it was handy having you here for that, but no I don't think id, I don't think I felt like like you were a teacher watching me or looking for the right answer or anything like that. So I don't think I was effected by it.		No, I mean I asked a couple of questions of you, and so I think it was handy having you here for that, but no I don't think id, I don't think I felt like like you were a teacher watching me or looking for the right answer or anything like that. So I don't think I was effected by it.			Um, the same, is that alright? No, I mean I asked a couple of questions of you, and so I think it was handy having you here for that, but no I don't think id, I don't think I felt like like you were a teacher watching me or looking for the right answer or anything like that. So I don't think I was effected by it.	

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
M4		I don't think much different.						Um, well at some point, I felt like I kind of trying to um, like um, really see what I am thinking. So, there was some kind of trying to that wouldn't have been there if I was alone. Um. Maybe slightly.
M5							And what about if I hadn't been say here listening, do you think the answers would have been the same? T Yep. Of course.	
M6		OK. So do you think if you hadn't been asked to think outloud do you think yout answers would have been the same or different? U Ohhh. Mostly the same. They'd be mostly the same. I don't think I did		And what about if I hadn't been sat her listening, do you think your answers would have been different? U Hm, yes, um, well maybe, because I asked you a couple of questions to clarify because I wasn't very sure				And what about if I hadn't been sat her listening, do you think your answers would have been different? U Hm, yes, um, well maybe, because I asked you a couple of questions to clarify because I wasn't very sure

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
		much wrong.		where about. So you being her was helpful for sure. Ye my answers might have changed a little. Especially when I asked you about, you know the pain killer question. Ye, you may have said to me flat out. No it wouldn't have made any difference, then it would have changed the outcome for sure.				where about. So you being her was helpful for sure. Ye my answers might have changed a little. Especially when I asked you about, you know the pain killer question. Ye, you may have said to me flat out. No it wouldn't have made any difference, then it would have changed the outcome for sure.
M7						Um. I think there is a potential that, um, the answers may have been different, um. That might maybe have an influence because you're in the room and obviously someone is listening to me. VB OK, that was the next question.		

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
						W Um, so. I'm not too sure if, if you weren't in the room. I would probably have just forgotten about the recorder. I don't know if potentially thinking out loud may have had an impact then or not. Um. So I'm not sure.		
M8		OK. Um. So, just thinking about the think aloud technique itself. Do you think that you would have given different answers if you hadn't been asked to think outloud? X I don't think so, I think I would have given the same answer.					OK, yep, and what about if I wasn't sat her listening, do you think your answers would be any different? X No, they would be the same.	
M9		OK, um. So, just					OK, um. So, just	

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
		thinking about the think aloud technique itself. Do you think if you hadn't been asked to think aloud. Do you think your answers would have been the same? Z Um, ye. Ye. VB And do you think if I hadn't been sat here listening do you think your answers would have been the same? Z Ye, ye, I don't think that would have made a difference to be honest. Ye. I'm not like, you know, its nothing like personal or something that you wouldn't want, you know. It doesn't really matter if someone is there I guess. Ye.					thinking about the think aloud technique itself. Do you think if you hadn't been asked to think aloud. Do you think your answers would have been the same? Z Um, ye. Ye.	

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
M10		I imagine that it would have been predominantly similar, but it might have been a different really if I wasn't thinking out loud. Um, I suppose because it is a bit of a kind of, you know I was trying to focus on what I was thinking and saying it out loud, rather than just having them, so it, so that kind of interferes with it, so you knew what I was thinking but, I don't know, I'm just looking at what I've put. Ye. I don't know, it probably would have been quite similar.					I don't think I was really bothered about that. I was just trying to think you know, what my thoughts were about the questions and stuff, so um.	
M11		They might be the same if I could focus as good as, I was					Um, I think it would be the same, ye, I think so, ye. I'm	

	9.2 Impact of TA on responses					9.3 Impact of interviewer listening		
	9.2.1 Uncertain	9.2.2 No impact	9.2.3 Changed responses			9.3.1 Uncertain	9.3.2 No impact	9.3.3 Changed responses
Transcript			9.2.4 More thoughtful if TA	9.2.5 Interviewer could correct me	9.2.6 Better constructed if not TA			
		thinking aloud, ye, I am pretty sure I mean I am 95% sure I would be the same.					trying to be as honest as possible, so ye.	