Doctors' Clinical Decision Making

Using Theory to Develop an Educational Intervention

Leila Mehdizadeh

The University of Leeds
School of Medicine

Submitted in accordance with the requirements for the degree of Ph.D.
Submitted for examination in July 2011

Acknowledgments

Many people have helped me to complete this doctoral thesis, all of whom I am indebted to. I would like to thank my supervisors Dr Hilary Bekker, Dr Naomi Quinton and Dr Vikram Jha for their support and guidance over the past four years. It has been a pleasure learning from them. The administrative staff of the following organisations helped me to collect data by kindly distributing the questionnaire to their members: Evidence Based Shared Decision Making, the Association for the Study of Medical Education, the European Association of Decision Making and the Society for Medical Decision Making. Dr Alan Schwartz accommodated the questionnaire at the 31st Annual meeting for the Society for Medical Decision Making and Dr Mark Graber generously offered to cover the administration fee to distribute the questionnaire electronically to the SMDM members. Postgraduate course tutors within the School of Medicine at the University of Leeds advertised my studies amongst their students. Also, Dr Matt Homer helped me with my SPSS queries.

The Islamic College, London provided me with a quiet workspace and people to talk to when I needed to get out of the house. My family and friends were encouraging and understanding when I did not attend social gatherings. Lastly, I am particularly grateful to my poor husband for all his patience, help and support. I am sorry that he has had to put up with the worst of me at the start of our married life!

Abstract

Medical education aims to train students to become safe and effective clinical practitioners. This includes the ability to make safe and effective clinical judgements and decisions (GMC 2009). It is assumed that trainee doctors acquire these skills through the hidden curriculum. This is not necessarily the case. There is reason to believe that medical education should include some explicit training for doctors to improve their clinical judgements and decisions. This is known as training in clinical reasoning. This thesis explored how to enhance doctors' clinical reasoning through effective training. The aims were to develop and evaluate an intervention informed by decision theory to improve doctors' reasoning about clinical judgements and decisions. A series of empirical studies were conducted to achieve these aims.

A systematic review and questionnaire study were conducted to evaluate existing interventions that aimed to enhance doctors' clinical reasoning skills. There was little agreement between medical educators on how to effectively enhance doctors' clinical reasoning through training. However, the minority of interventions that aimed to improve doctors' awareness about their own reasoning processes were effective.

Little is known about how to improve the processes doctors use to make clinical judgements and decisions in practice. A qualitative interview study was conducted to explore doctors' views and experiences of how to make effective clinical judgements and decisions. Doctors had limited explicit insight into their own

reasoning processes, such as the methods that lead to good decisions and factors that bias their reasoning.

A quasi-experimental study was developed to evaluate the feasibility of an intervention to enhance doctors' understanding about their own reasoning processes. A brief tutorial was shown that explained the basic science underpinning human judgement and decision making. Doctors were receptive to learning about this information. They found it relevant to their clinical practice and gained knowledge about decision sciences concepts. Findings from this thesis suggest that, potentially, doctors can improve their clinical judgements and decisions through training to understand how they think about clinical problems.

Contents

Contents in Full

Tables

Key for abbreviations used

ASSIA	Applied Social Sciences Index and Abstracts
BOS	Bristol Online Surveys
CDSS	Computer Decision Support System
CPD	Continuing Professional Development
CRD	Centre for Reviews and Dissemination
EDREC	Medicine and Dentistry Educational Research Ethics Committee
ERIC	Educational Research Information Centre
EPPI centre	Evidence for Policy and Practice Information and co-ordinating centre
FY1/FY2	Foundation Year 1 / Foundation Year 2
GMC	General Medical Council
GP	General Practitioner
MCQ	Multiple-Choice Question
MRC	Medical Research Council
NHS	National Health Service
RCT	Randomised Controlled Trial
SMDM	Society for Medical Decision Making
SN	Study Number
SPICES	Student centred, problem based, integrated teaching, community based, electives, systematic
SPSS	Statistical Package for the Social Sciences
UK	United Kingdom
USA	United States of America

Presentations of relevance

MEHDIZADEH, L., BEKKER, H. L., JHA, V., and QUINTON, N. D. (2008) Interventions to facilitate clinical decision making: a systematic review. Poster presentation at the Leeds Institute of Health Sciences Symposium, Leeds, 30 June 2008.

——— (2008) Interventions to facilitate clinical decision making: a systematic review. Oral presentation at the Medical Education Research Development: 1st Annual Yorkshire Regional Meeting, Leeds, 10 July 2008.

——— (2008) Interventions to facilitate clinical decision making: a study protocol. Poster presentation at the annual conference for the Faculty of Medicine and Health, University of Leeds, Leeds, 12 November 2008.

——— (2009) Doctors experiences of making clinical decisions effectively: an interview study. Oral presentation at the Medical Education Research Development: 2nd Annual Yorkshire Regional Meeting, Leeds, 25 June 2009.

——— (2009) Doctors experiences of making clinical decisions effectively: an interview study. Oral presentation at the Leeds Institute of Health Sciences symposium, Leeds, 29 June 2009.

——— (2009) Training to improve doctors' decision making: a systematic review. Poster presentation at the 31st annual meeting for the Society of Medical Decision Making, Hollywood, USA, 18-21 October 2009.

——— (2010) Training to improve doctors' decision making: a systematic review. Oral presentation at the 5th International workshop on clinical reasoning. Kings College London, 17-18 April 2010.

——— (2010) Doctors' views about making clinical decisions effectively: an interview study. Oral presentation at Bradford Royal Infirmary, Maternity Unit, Bradford, 7 July 2010.

——— (2010). Doctors' views about making clinical decisions effectively: an interview study. Poster presentation at the Association for the Study of Medical Education (ASME) Annual Scientific Meeting, Cambridge, 21-23 July 2010.

——— (2010) Doctors' views about making clinical decisions effectively: an interview study. Poster presentation at the Association for Medical Education in Europe (AMEE) annual conference, Glasgow, 4-8 September 2010.

1.

Introduction: concepts, theory and evidence of training doctors in clinical reasoning

This chapter provides an overview of issues related to training doctors to make better clinical judgements and decisions. It starts with a discussion on the importance of good judgements and decisions in medicine. It introduces the theories and evidence about how people in general and doctors specifically make judgements and decisions. It provides evidence from a review of the literature about existing work on clinical reasoning training. The challenges and broader issues that influence the development of effective training are considered. The final part of the chapter identifies the aims and objectives of the thesis that are addressed in subsequent chapters.

1.1. Importance of clinical judgements and decisions in practice

It has been argued that a doctor spends more time making judgements and decisions about clinical cases than any other work related activity (Croskerry 2000). The ability to make good judgements and choices underlies the entire care process; from gathering clinical information, interpreting and synthesising data, making accurate and timely diagnosis and administering appropriate treatment plans (GMC 2009). In the clinical context, making good judgements and decisions is an acquired skill. Doctors must be able to make judgements and decisions well in risky and/or uncertain situations. In risky situations, doctors make choices when the probabilities of an outcome are known to the doctor (Wu, Zhang, and Gonzalez 2004; Baron 2008). An example is deciding whether to perform a particular operation on a patient or not. In uncertain situations, a doctor does not know the probabilities of an outcome and must estimate them to make a choice (Wu et al. 2004). Doctors also face working under emotive, time pressured situations with limited resources. These factors can also impact on their judgements and decisions, and the risk of making serious cognitive errors can be greater under such circumstances (Means, Salas, Crandall, and Jacobs 1993; Croskerry 2003).

1.2. Error in medicine

Most of the work on error in medicine has focused on mistakes related to diagnosis. Diagnostic errors are associated with higher patient morbidity than other types of medical errors (Brennan, Leape, Laird, Hebert, Localio, Lawthers, Newhouse, Weiler and Hiatt 1991; Wilson, Runciman, Gibberd, Harrison, Newby and Hamilton 1995). Evidence from autopsy studies suggests that the risk of making a diagnostic error is between 10-15% (Goldman, Sayson, Robbins, Cohn, Bettmann and Wesiberg 1983; Kirch & Schafii 1996; Shojania, Burton, McDonald, and Goldman 2003; . It has been found that most diagnostic errors are caused by incorrect thinking about a clinical case. In other words, errors of a cognitive nature are more common that other types of medical error (Clarke, Spejewsk, Gertner, Webber, Hayward, Santora, Wagner, Baker, Champion, Fabian, Lewis, Moore, Weigelt, Eastman and Blank-Reid 2000; Kuhn 2002; Graber, Franklin, and Gordon 2005). These findings are significant as they demonstrate a doctor's vulnerability to make poor judgements and decisions in practice. While it is unlikely all medical error can be eradicated entirely, some authors believe that doctors' cognitive errors are preventable (Clarke et al. 2000; Famularo, Salvini, Terranova, and Gerace 2000; Glick, Workman, and Gaufberg 2000; Kuhn 2002; Croskerry 2002; Croskerry 2003). The introduction of formal training in medical education to improve doctors' clinical judgements and decisions is one possible way to reduce cognitive errors. This is often referred to as training in clinical reasoning (Higgs & Jones 1995). If this type of training is successful, the implications are an increase in patient safety and quality of care, cost-effective use of resources and a reduction in complaints or legal action against doctors.

1.3. Introducing clinical reasoning training into medical education

Medical education aims to train students to become safe and effective clinical practitioners. This includes the ability to make clinical judgements and decisions safely and effectively (GMC 2009). However, the development of judgement and decision making skills has received insufficient attention in medical education. Doctors do not receive training in clinical reasoning as a formal part of their medical education. There is an implicit assumption that doctors acquire good clinical reasoning through the hidden curriculum, usually by opportunistic observation of reasoning demonstrated by expert doctors (Howe, Holmes, and Elstein 1984; Chessare & Lieu 1998; Round 1999). The limitation with this approach is that their strategies and habits may be copied with little consideration to their value (Chessare & Lieu 1998).

There is debate in the literature about the value of introducing formal training in clinical reasoning into medical education. One view is that doctors can and should receive training that will improve their clinical judgements and decisions (Elstein, Shulman, and Sprafka 1978; Elstein 1981; Elstein, Rovner, and Rothert 1982; Kassirer 1989; Croskerry 2000; Graber et al. 2005; Croskerry 2009a). For more than thirty years, researchers have sought to identify and understand the reasoning processes doctors use to make judgements and decisions. This line of work was driven by the assumption that the reasoning processes of expert doctors could be taught to inexperienced doctors. This approach has been criticised as it assumes that by identifying these reasoning processes, doctors can be taught the generic underlying principles of making good judgements and decisions.

An alternative view is that there are no general principles that doctors can learn to enhance their judgements and decisions. Rather, the ability to make good judgements and decisions is acquired through years of repeated practical experience (Norman 2005). This view also rests on the assumptions that there is no one ideal way to solve a clinical problem, and the nature of learning is context specific (Eva, Neville, and Norman 1998; Schuwirth 2002; Nendaz & Bordage 2002; Norman, Eva, and Schmidt 2005; Eva & Norman 2005; Norman & Eva 2010). This perspective implies that it is impractical to try and teach general principles or strategies of good judgement and decision making when evidence suggests that people do not recognise how to apply their learning to novel situations in different contexts (Gick & Hollyoak 1983; Norman, Tugwell, Feightner, Muzzin, and Jacoby 1995).

1.4. Challenges of enhancing doctors' clinical reasoning through formal training

Understanding how to improve doctors' clinical judgements and decisions through formal training is challenging. This is partly because terminology is poorly defined in the clinical reasoning literature and knowledge of several disciplines is required in order to understand how effective training programmes can be developed. These issues are discussed further in the sections below.

1.4.1. Defining terms from the decision sciences and clinical reasoning literature

In the decision sciences, judgement is a separate area of study from decision making. Their theoretical underpinnings and definitions are distinct. Judgement can be distinguished from decision making as the assessment of alternatives to form an opinion. Whereas making a decision can be defined as a choice of action or inaction between alternative options (Dowie 1993; Baron 2008). However, this distinction cannot be applied to the clinical judgements and decisions of doctors. Making a clinical decision is a more difficult concept to define because it does not just involve choosing a course of action for a patient. The doctor must gather the relevant clinical information through questioning, examining the patient and possibly through clinical investigations. This information then needs to be interpreted to understand the problem before a diagnosis can be made. Hereafter the doctor decides how to manage the condition, which could include active treatment, referral to a specialist, a period of watchful waiting or discharge (Croskerry 2002). In practice it is not clear when clinical judgements become clinical decisions. The clinical reasoning literature does not make a distinction between a clinical judgement and clinical decision. Rather, clinical decision making is often used synonymously with clinical judgement, clinical inference, clinical reasoning, diagnostic reasoning and clinical problem solving (Thompson & Dowding 2002; Norman 2005).

The study of doctors' clinical judgements and decisions are both part of an area of research known as clinical reasoning (Higgs & Jones 1995). This thesis was concerned with how to enhance both the clinical judgements and decisions of doctors. For this reason, the term clinical reasoning was used to refer to the cognitive processes informing both clinical judgements and decisions. This thesis was concerned with training doctors to improve their clinical reasoning. However, given that it is more correct from a decision sciences perspective to distinguish judgements from decisions, it was helpful on occasion to separate clinical judgements from clinical decisions in the following way. Making a judgement in the clinical context and improving accuracy of judgements is associated with forming an opinion about a diagnosis. Whereas a clinical decision is associated with making a choice about how to manage a patient's condition. In other words, doctors make judgements to diagnose a condition and decisions to manage the condition. At some points in the thesis this distinction has been used when referring to either clinical judgements specifically or clinical decisions. Below is a table that summarises how these key terms have been used in this thesis.

Decision sciences context	Definition	Clinical context	Definition
Judgement	Assessment of alternatives to form an opinion	Clinical judgement	Forming opinions about diagnosis
Decision	Choice of action or inaction between alternative options	Clinical decision	Choosing how to manage patient's condition
		Clinical reasoning	Cognitive processes that inform both clinical judgements and decisions

Table 1: Defining judgements and decisions.

Most of the clinical reasoning literature is about understanding the diagnostic process and enhancing diagnostic accuracy (Norman 2005). Further, when the term clinical decision is used in the literature, it is almost always referring to a diagnosis. Much less is known about the processes doctors use to make patient management decisions or how to enhance the way they make treatment choices (Norman 2005). A lot of work has been conducted on improving treatment decision making, but this work aims to help patients make better choices about their health. Interventions that have been designed to enhance treatment decision making include patient decision aids and ways to encourage shared decision making between the patient and doctor (O'Connor, Rostom, Fiset, Tetroe, Entwistle, Llewellyn-Thomas, Holmes-Rovner, Barry, and Jones 1999; Bekker 2010; Légaré, Ratté, Stacey, Kryworuchko, Gravel, Graham, and Turcotte 2010). These areas are separate from the clinical reasoning literature and are not referred to in detail in this thesis.

In order to help doctors make better clinical judgements and decisions, there needs to be an understanding of what makes a judgement or decision good. It is debated whether judgement and decision quality should be based on outcomes (accuracy of diagnosis, success of treatment) or processes such as making a well reasoned choice (Fischhoff 2002; Hammond 2007; Bekker 2010). The risky and uncertain nature of medicine described in Section 1.1, means that poor judgements and decisions can result in good medical outcomes and good judgements and decisions can result in adverse medical outcomes. This is summarised in Table 2.

	Good judgement/decision	Poor judgement/decision
Good medical outcome	✓✓	✗
Adverse medical outcome	✓	✗✗

Table 2: Defining quality of judgements and decisions.

The most desirable situation is when doctors make good judgements and decisions and the medical outcomes are also good. However it is inevitable that some patients will not respond well to treatments or ever recover from an illness or injury. In situations where the medical outcome is adverse, doctors' judgements and decisions may be challenged by colleagues and patients. They must be able to justify that the steps taken to reach their decisions were acceptable, even though a patient's condition worsened (GMC 2009). If it transpires that the process was good, then it does not necessarily mean that the doctor made a misjudgement or poor decision. It is a concern when doctors make poor judgements and decisions, even if the medical outcome is good. Poor decision making compromises patient safety and doctors risk facing litigation if their poor practice leads to adverse medical outcomes (Croskerry 2003). For these reasons, an increasing number of decision making researchers stress that the reasoning process is a better marker of decision quality than the consequence of the decision (Frisch & Clemen 1994; Sox 1999; Yates, Veinott, and Patalano 2003; Baron 2008; Schwartz & Bergus 2008; Bekker 2010). This definition of decision quality is particularly relevant in medicine as it is a requirement for doctors to be able to always justify their decisions and actions (GMC 2009).

In this thesis, a good judgement or decision is defined as one that is reasoned well. However, using this definition presented a challenge in terms of measuring a good reasoning process as there is no consensus on what constitutes good and poor thinking (Bekker 2010). Previous measures that have been used to indicate good clinical reasoning processes were critical thinking about clinical cases (Round 1999; Abraham, Upadhya, Torke, and Ramnarayan 2004; Beullens, Struyf, and Van Damme 2006) and evidence of a more considered use of heuristics and avoidance of bias (Wolf, Gruppen, and Billi 1988; Hershberger, Part, Markert, Cohen, and Finger 1995). There is no single measure that is widely accepted as a good measure of clinical reasoning quality.

1.4.2. An interdisciplinary approach to clinical reasoning training

Three different areas of literature were drawn on in order to inform understanding of how to improve doctors' clinical reasoning. In general, training interventions should be designed with reference to a theoretical basis so that underlying assumptions are explicit and possible mechanisms that contribute to the effectiveness of interventions can be identified (Chen 1990; MRC 2008). The decision sciences are the field of study about human judgement and decision making. They draw on theories, methodologies and evidence from the behavioural sciences, economics and mathematics to describe how all individuals make judgements and decisions. Evidence from the decision sciences provides a context for understanding that all people make judgements and decisions in largely the same way. Knowledge about the theories and evidence that describe how doctors, specifically, make judgements and decisions in the clinical context was also necessary. Finally, knowledge about medical education was necessary. For instance, it was important to have an understanding of the different methods that are currently used to train and assess doctors' competencies. Theoretical perspectives and the medical education literature are discussed further in the following sections.

1.5. Theoretical perspectives of judgement and decision making

There are two main types of theories that can inform the training of doctors in clinical reasoning. These are theories about generic and expert decision making. Generic theories explain the processes that all individuals use to make everyday judgements and decisions. Expert theories explain the processes that doctors use to make clinical judgements and decisions. There are many generic and expert decision theories, but only a few are discussed in further detail below. These include the information processing approach, as an example of a generic decision theory and two expert theories of clinical judgement.

1.5.1. Understanding how all individuals make judgements and decisions: an information processing approach

The information processing approach has driven much of the research on judgement and choice (Payne & Bettman 2004). It is still one of the dominant frameworks to describe how all individuals make judgements and decisions from the information 'out there' (Payne & Bettman 2004; Baron 2008; Croskerry 2009a). According to the information processing approach, the brain has a given infrastructure that makes sense of and stores information from the outside world (Newell & Simon 1972). People need to process information when solving problems and making decisions, but the brain cannot consciously process all of the complex information 'out there'. The brain has a limited capacity for conscious attention (Simon 1972). Consequently, people are highly selective about the information upon which they focus. Their judgements and decisions are based on information which has been processed and not the full information available. The mechanisms people use to solve complex problems with a limited processing capacity are known as heuristics (Payne & Bettman 2004). These are simple rules of thumb that help people take short cuts when making judgements and decisions.

Heuristics are generally accepted as one of the two main strategies that all individuals use to process information. The other strategy is known as a systematic strategy. Together heuristic and systematic strategies represent two systems of information processing. The way people use these two systems is represented in dual processing models of information processing (Chaiken 1980; Chen & Chaiken 1999; Payne & Bettman 2004). A dual processing model proposes a system 1 and system 2 mode of thinking (Stanovich & West 2000). System 1 thinking is characterised by a fast and frugal process via the use of mental shortcuts i.e. heuristics. Judgements and decisions are made quickly and easily because little, if any, conscious thought is involved. For example, people readily conclude that a baby dressed in pink is a girl, or they may buy a camera that is recommended by a trusted friend rather than spend time searching for the latest model or best value for money. Heuristics usually lead to correct judgements and satisfactory decisions, but this is not always the case. Occasionally, people can unknowingly focus on irrelevant information and ignore relevant information (Payne & Bettman 2004). This can lead to inaccurate judgements and poor choices. For example, choosing to attend a university based on the number of friends who also plan to go to the same university, rather than prioritising course quality, reputation of the university, fees and living costs etc.

System two thinking is characterised by a systematic strategy. This is a more complicated and deliberative process that feels challenging and takes time to make a decision. It requires conscious effort to evaluate the advantages and disadvantages of the consequences of several options. Systematic processing is sometimes used to reach decisions that have important consequences, for example when purchasing a house. Whilst this approach requires more effort to reach a decision, it is more reliable, errors are unlikely, people are less likely to regret their decisions and can justify them to others when necessary.

Doctors use the same infrastructure to make judgements and decisions as everyone else. Some of their judgements and decisions are made using a heuristic method and at other times a systematic method may be used. For example, doctors are taught to always consider worst case scenarios in some cases such as pre-eclampsia when a pregnant woman feels breathless or angina in a person that experiences chest pain at rest. Doctors may also follow the advice of a more experienced colleague when unsure about the type of treatment to give a patient. These are examples of heuristic methods because they simplify the way a clinical problem is solved. Doctors will think systematically about most novel clinical problems or when the situation involves high risk to the patient's health. For example, if a patient is not responding to a treatment in the way that is expected, a doctor may revisit the details of the case and consider alternative diagnoses. To decide whether to perform surgery that involves significant risk to a patient's health, the doctor may weigh up the risks and benefits of performing or not performing the surgery. These are examples of systematic methods because conscious effort is made to deliberate about the clinical problem.

1.5.2. Expert clinical judgement theories: hypothetico-deductive and pattern recognition methods

A number of theories have been proposed to describe how doctors make clinical judgements about diagnosis. However, it is the hypothetico-deductive and pattern recognition methods that have been the most influential in furthering understanding about the diagnostic process. These are outlined briefly below.

1.5.2.1. Hypothetico-Deductive method

The Hypothetico-Deductive method is one of the earliest theories of clinical diagnosis (Elstein, Kagan, Shulman, Jason, and Loupe 1972). It describes diagnosis as a process of hypothesis generation using deductive reasoning i.e. the conclusion necessarily follows the hypothesis (Eysenck & Keane 2000). Early observational studies established that, within a few seconds or minutes of first seeing a patient, the doctor considers one or more hypotheses (Elstein et al. 1978; Barrows, Norman, Neufeld, and Feightner 1982). These hypotheses are used to guide subsequent data collection, for instance, searching for other signs and symptoms that are expected to be present in a particular condition (Elstein & Schwartz 2002). Each hypothesis is tested for how well it fits the medical condition and the doctor eliminates those hypotheses that are no longer accurate. The hypothesis that is deemed most accurate after the testing stage is accepted as the diagnosis. Even though generating and testing hypotheses has been described as a serial process, in reality the diagnostician performs these simultaneously and in a rapid, automatic way (Elstein et al. 1978).

The Hypothetico-Deductive theory has received some empirical support (Barrows & Tamblyn 1980; Donnelly, Sisson, and Woolliscroft 1990; Nakamura 2008) however it was soon challenged on the basis that both expert and novice diagnostician were reportedly using a similar hypothesis testing strategy. This finding was at odds with the view that expert and novice doctors do not solve clinical problems in the same way (Patel & Groen 1986; Elstein & Schwartz 2002; Norman 2005). Evidence shows that expert doctors do not usually make a diagnosis by generating and testing a set of hypotheses. It was found that they were able to make accurate diagnoses quicker and more efficiently using a recognition strategy (Groen & Patel 1985; Schmidt, Norman, and Boshuizen 1990; Brooks, Norman, and Allen 1991; Norman, Coblentz, Brooks, and Babcook 1992).

1.5.2.2. Pattern Recognition method

An alternative theory proposed was a Pattern Recognition method, that described diagnosis as a categorisation process (Kulikowski 1970; Groen & Patel 1985). The doctor categorises the signs and symptoms of medical conditions and stores these categories in clinical memory (Higgs & Jones 1995). When a new case is presented to the doctor, any similar signs and symptoms are recognised and placed into an existing category and given the same diagnostic label (Higgs & Jones 1995). The already existing categories serve as a prototype that enables doctors to recognise and match clinical patterns to existing diagnostic categories. Overall, this method describes an ability to immediately recognise many medical conditions based on their collection of their signs and symptoms. In contrast with the hypothetico-deductive method, pattern recognition involves inductive reasoning and is also known as a bottom-up theory. This means that information processing is directly influenced by environmental stimuli, i.e. the signs and symptoms of a medical condition and that specific information is examined to arrive at a general conclusion (i.e. a 'diagnosis') (Eysenck & Keane 2000).

Pattern recognition has also received empirical support (Groen & Patel 1985; Schmidt et al. 1990; Brooks et al. 1991; Norman et al. 1992). Further, other evidence found pattern recognition was most associated with diagnostic success when compared to other reasoning strategies (Patel & Groen 1986; Coderre, Mandin, Harasym, and Fick 2003). A limitation of the pattern recognition theory is that it rests on the premise that a doctor must have previously experienced and categorised a particular clinical pattern in order to recognise any similar patterns. It offers no explanation as to how doctors are able to diagnose conditions of which they have no previous experience. In other words how the initial prototype patterns develop is unaddressed. The Hypothetic-Deductive and Pattern Recognition theories propose two different methods of clinical diagnosis. They have since been reconciled in the following way; in difficult or novel clinical situations doctors may use a hypothetico-deductive method otherwise diagnosis is usually a direct and automatic process of pattern recognition (Elstein & Schwartz 2002).

1.5.3. Theoretical context for improving doctors' clinical reasoning

When evaluating the suitability of the hypothetico-deductive and pattern recognition method to inform a training intervention in clinical reasoning, the following limitations were noted. First, they did not offer guidance on how to enhance the quality of judgements and decisions. For example, there was an explanation of how incorrect judgements occur using both of these theories, but this information alone is of little use if there is no understanding of how to avoid errors. Second, doctors can make a diagnosis by drawing on other methods. In brief these include, ruling out the worst case scenario, an exhaustive search for and evaluation of all the medical facts, trusting intuitive feelings, following established clinical guidelines, and using established statistical proofs to calculate accurate probabilities of disease (Croskerry 2002; Croskerry 2009b). This suggests doctors utilise numerous reasoning styles to solve clinical problems depending on level of expertise and the clinical presentation (Elstein & Schwartz 2002). Evidence shows that it is better that doctors do not limit themselves to one mode of thinking. Greater diagnostic success has been associated with explicit instruction to use a combined reasoning approach of both heuristic and systematic methods compared to when no instruction was given (Ark, Brooks, and Eva 2006; Ark, Brooks, and Eva 2007; Eva, Hatala, LeBlanc, and Brooks 2007).

1.5.4. An integrated model of clinical reasoning

Recently, a framework has been proposed that brings together generic theory of judgement and decision making from the decisions sciences with expert theory of diagnosis (Croskerry 2009b). Drawing on the established division between heuristic and systematic thinking, the various processes used to reach diagnosis can be grouped under two systems. Intuitive, heuristic approaches that involve no deliberate thought such as pattern recognition, gut instinct and learned rules of thumb are grouped under system 1. Whereas conscious, systematic approaches such as the hypothetico-deductive method or calculating disease probabilities are grouped under system 2 (Croskerry 2009b; Norman 2009). Figure 1 illustrates this dual processing model of diagnostic reasoning. It models the doctor's mind in the following way: the presentation of the patient's condition is either recognised or not by the doctor. At the simplest level, if it is recognised the system 1 processes are engaged immediately and automatically to identify the cause of the condition. This is considered to be the default method of diagnosis. If the presentation is not recognised, then the slower systematic processes of system 2 are engaged to make a diagnosis (Croskerry 2009b). It is proposed that system 1 processes represent a set of reflex systems that have adapted for survival, while system 2 processes represent the logical, rational part of the brain that develop through learning (Croskerry 2009b).

The two systems are not mutually exclusive from each other. While an unfamiliar clinical presentation is usually diagnosed using a system 2 process, repeated exposure leads to recognition and eventually a system 1 process is engaged. Furthermore, it is possible that a doctor may initially be engaged in one system of thinking before switching to the other. For example an initial assessment of a patient with vomiting and abdominal pain may suggest a gastroenteritis diagnosis, but if the patient's condition does not resolve in the expected manner system 2 can force a reassessment of the diagnosis. System 2 can function like a monitor of system 1 in order that the doctor stops to reflect critically on clinical situations when needed. This is referred to as a rational override (Croskerry 2009b). Alternatively, a doctor may be aware that applying an established statistical proof like Bayes theorem can help estimate the probability of a particular diagnosis more accurately, but chooses to rely on intuitive feelings in the actual clinical context. This describes an irrational override of system 2 by system 1.

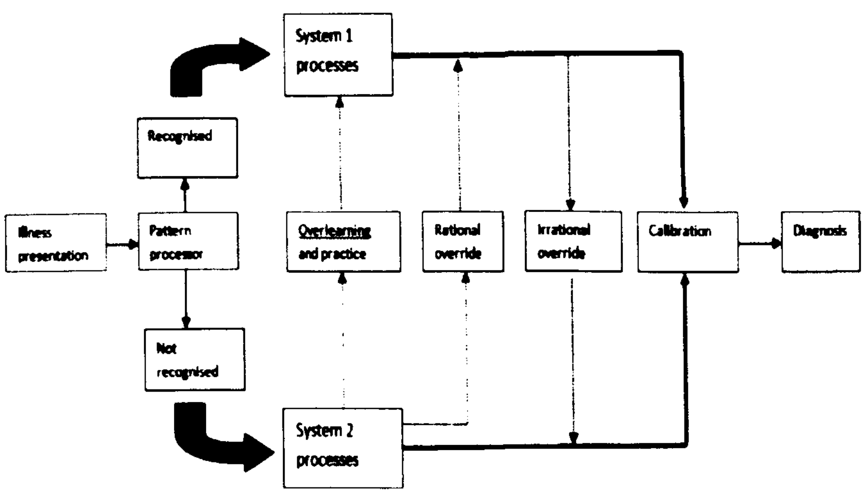


Figure 1: Dual processing model of doctors' diagnostic reasoning (adapted from (Croskerry, 2009b).

A limitation of the dual processing model is that it only describes the processes that underpin diagnosis. It does not describe the way doctors make decisions about managing patients. Therefore the framework could offer a useful theoretical basis for intervening in doctors' diagnostic judgements, but not for treatment choices. Its strengths are (i) a coherent translation of the decision sciences in a way that may be of use for the busy doctor (ii) the doctor is positioned in a broader context of human cognition to demonstrate they apply generic as well as expert reasoning processes in the clinical context, (iii) it reflects the complexity of clinical reasoning by integrating the repertoire of approaches used to make a diagnosis (iv) it shows how errors can arise occur in doctors' thinking e.g. not recognising an atypical presentation (v) it suggests how the quality of clinical judgements might be improved, for instance a cautious use of system 1 processes and/or developing critical thinking skills to strengthen the use of system 2 processes.

1.6. Evidence of heuristics and bias in clinical judgement and decision making

Section 1.5.1 introduced the idea that people often simplify the way they make judgements and decisions by employing rules of thumb, known as heuristics. A large body of research has been devoted to identifying the heuristics people use and exploring the effects they have on human judgement and decision making. Most of this work is attributed to an extensive research programme by Kahneman and Tversky in the 1970's, where many types of heuristics and their characteristics were identified. Heuristics are recognised as a universal feature of human cognition (Croskerry 2003; Norman 2009). They are a resourceful, practical and effective way of achieving quick and satisfactory judgements and decisions (Croskerry 2000). Most of the time they will serve people well but occasionally a price will be paid for taking shortcuts (Croskerry 2000). A key finding is that heuristics are associated with systematic and predictable errors (Kahneman, Slovic, and Tversky 1982). That is, people make the same types of incorrect judgements when a particular heuristic strategy is used (Kahneman & Tversky 1972; Kahneman & Tversky 1973; Tversky & Kahneman 1973). These cognitive errors are referred to as biases.

A number of these biases have been illustrated in the clinical context (Chapman & Elstein 2000; Bornstein & Emler 2001; Croskerry 2002). This does not mean that doctors are inherently bad decision makers, but they are human and have the same infrastructure as others to use when making clinical judgements and decisions (Elstein 1999). Biased information processing can occur in the clinical context at any stage of the doctor-patient interaction with potentially grave consequences. There are a large number of heuristics that have been identified that can be grouped under three main types: (i) confirmation, (ii) anchoring and (iii) availability (Kahneman et al. 1982). The following paragraph offers a brief description of how doctors may use each of these heuristics and the biases associated with their use.

The confirmation heuristic is used to seek only data that confirms a preconceived diagnostic hypothesis, rather than disconfirm it. Disconfirming evidence is undervalued such as the absence of particular symptoms. The biases associated with this heuristic can be a fixation on a weak hypothesis that is based on first impressions rather than appropriate evidence. Failure to test other hypotheses may result in wasted time and effort and a doctor risks missing a correct diagnosis (Kahneman et al. 1982; Joseph & Patel 1990; Elstein 1999; Chapman & Elstein 2000; Croskerry 2002; Croskerry 2003; Baron 2008; Norman 2009). The Anchoring heuristic involves the fixation on salient features of a presentation early in the diagnostic process. This can lead to the failure to adjust initial impressions in light of later information, and therefore the correct diagnosis may be missed (Kahneman et al. 1982; Croskerry 2002; Croskerry 2003; Norman 2009). The availability heuristic is used when a diagnosis is judged to be more likely if it is easily retrievable from memory. Common conditions are readily considered, whilst conditions that are uncommon less readily come to mind. This may lead to inaccurate estimates of disease probabilities, and therefore inaccurate diagnosis (Kahneman et al. 1982; Chapman & Elstein

2000; Bornstein & Emler 2001; Croskerry 2002; Croskerry 2003; Baron 2008; Norman 2009).

Doctors, like other people, use heuristics automatically and unconsciously. This means that unknowingly, they risk making incorrect judgements and poor decisions. In order to reduce the risk of these cognitive errors, it has been suggested that doctors should receive training in cognitive debiasing (Kassirer & Kopelman 1991; Croskerry 2000; Croskerry 2002). Cognitive debiasing involves learning to recognise the situations in which cognitive errors can arise and knowing how to avoid them (Croskerry 2002; Croskerry 2003). A number of debiasing techniques have been suggested by (Croskerry 2003) but their effectiveness for training purposes remains largely untested. There is evidence that raising doctors' awareness of heuristics and bias can improve aspects of their clinical reasoning on hypothetical scenarios (Wolf et al. 1988; Gruppen, Margolin, Wisdom, and Grum 1994; Hershberger et al. 1995; Round 1999; Abraham et al. 2004).

An encouraging find in one study was reported by Bornstein, Christine Emler, and Chapman (1999). These researchers found that doctors were not influenced by the sunk-cost effect in medical situations, but were in non-medical situations. The sunk cost-effect describes a phenomenon where people are reluctant to consider alternative options because they have invested a lot of time, effort and/or money into a particular decision (Croskerry 2003). Treatment decisions were not biased by the amount of time and money already invested in a diagnosis and particular treatment plan, but these factors did bias non-medical decisions (Bornstein et al. 1999). Despite the extensive work that has been carried out in this area, there is little evidence to suggest that gaining knowledge about heuristics and bias would improve actual judgements and decisions in practice. In the non-clinical context, awareness training about heuristics was not enough to improve their judgements in practice (Fischhoff 1975; Welsh, Begg, and Bratvold 2007).

1.7. Evidence of training doctors in clinical reasoning

The interest in formally training doctors to make better judgements and decisions began in the 1970s with The Medical Inquiry Project (Elstein et al. 1978). Elstein and colleagues were motivated to understand two main issues, namely, (i) how doctors reach a clinical diagnosis, and (ii) how to help doctors improve their diagnostic skills. From this work, it was proposed that diagnosis was a process of generating and testing multiple hypotheses (see Section 1.5.2.1). As a first step in developing a formal training course to improve clinical reasoning, Elstein and colleagues conducted two surveys on members of the Society for Medical Decision Making. Here they identified the individuals and institutions that were offering clinical reasoning courses for doctors, the types of courses that existed and obtained their views and preferences towards future courses (Elstein 1981; Elstein, Dawson-Saunders, and Belzer 1985). The vast majority of training was happening in North America. In 1981, twenty different institutes offered formal courses to doctors in clinical reasoning, and twelve offered a course that included some content on judgement and/or decision making (Elstein 1981). A few years later, sixty institutes were including some formal training to clinical reasoning (Elstein et al. 1985). It was not clear how many of these courses were devoted entirely to improving doctors' clinical reasoning or whether they had been integrated fully into medical curricula.

The survey conducted by Elstein and colleagues found that the majority of courses aimed to teach doctors how to apply normative decision theories to their clinical judgements and decisions. Normative decision theories describe how people should make decisions if they want to select the optimum or rational choice (Baron 2008). Decision making from this approach is based on mathematical and statistical proofs that require people to calculate the optimum choice. Most of the clinical reasoning courses taught doctors practical ways of applying Bayes theorem to clinical diagnosis and Expected Utility Theory (EUT) to treatment choices. Bayes theorem consists of a statistical formula that can be applied to the clinical context as a way of helping doctors make accurate estimates of disease probabilities (Moreira, Bisoffi, Narvaez, and Van den Ende 2008). The principles of EUT are applied through decision analysis methods. Decision analysis offers ways of graphically representing decision problems with all of the options and consequences explicitly represented (Grant, Keim, and Telfer 2006). Most of the courses that were surveyed taught doctors how to make optimal treatment choices by evaluating the possible consequences of each treatment option against their own preferences (van der Velde 2005; Moreira et al. 2008).

These findings indicate that medical educators prioritise a statistical approach to inform the design of clinical reasoning training where the focus is on improving the outcomes (e.g. accuracy) of doctors' judgements and decisions (see Section 1.4.1). Other types of clinical reasoning courses were identified in the literature that also takes this approach. Examples include encouraging the use of guidelines and computer decision support systems that enhance diagnostic accuracy (de Dombal, Dallos, and McAdam 1991; Lagerlov, Loeb, Andrew, and Hjortdahl 2000; Akici, Kalaça, Ugurlu, Karaalp, and Çali S 2003; Hedrick & Young 2008). Also, manikins and standardised patients have been used to allow doctors to practice making real decisions in a safe learning environment (Chopra, Gesink, De Jong, Bovill, Spierdijk, and Brand 1994; Byrne, Sellen, Jones, Aitkenhead, Hussain, Gilder, Smith, and Ribes 2002).

An alternative approach to improving doctors' clinical reasoning has been identified in the clinical reasoning literature. This approach rests on the view that it is more important to teach doctors how to reason well about their judgements and choices. Applying normative decision theory to clinical problems is cumbersome, time consuming and an impractical way of enhancing the busy doctor's judgements and decisions in practice (Croskerry 2005). Improving doctors' ability to reason better about clinical problems has not received sufficient attention in medical education. This approach requires a translation of the work from the decision sciences, which are mostly unknown to doctors (Schwartz & Bergus 2008). Some authors argue that doctors may gain benefit from an understanding of system 1 and system 2 processes (see Section 1.5.1), the factors that contribute to poor thinking and ways to avoid them, and the importance of monitoring one's own thought processes (Croskerry 2002; Graber, Gordon, and Franklin 2002; Croskerry 2003; Schwartz & Bergus 2008). In this thesis, the term decision literacy has been introduced to describe an explicit awareness of, and ability to critically evaluate, one's own reasoning processes. There is little evidence that the effectiveness of enhancing doctors' decision literacy has been evaluated.

1.8. Broader issues in developing a training programme in clinical reasoning

A clinical reasoning course for doctors has to be designed, implemented and evaluated in the context of medical education (Campbell, Murray, Darbyshire, Emery, Farmer, Griffiths, Guthrie, Lester, Wilson, and Kinmonth 2007). This requires an understanding of the graduate outcomes as required from governing medical bodies, and current trends in learning and medical curriculum design (Barrow, McKimm, and Samarasekera 2010). To optimise the effectiveness of a new training course, the teaching, learning and assessment methods should reflect current good practices in medical education (Barrow et al. 2010). These issues are discussed in the following sections.

1.8.1. Trends in medical education

Medical education, like other disciplines, has gone through trends and shifts in philosophy in response to workplace demands and advances in learning (McKimm 2010). The emphasis has moved away from a teacher centred towards a learner centred curriculum. This involved a decrease in traditional lecture based teaching in favour of a more interactive problem based learning approach, an integrated interdisciplinary curriculum and an emphasis on student self-directed learning (Barrow et al. 2010). These recommendations were based on a model of curriculum developed, SPICES model proposed by Harden, Sowden, and Dunn (1984). SPICES is an acronym that stands for student centred, problem based learning, integrated teaching, community based, electives and systematic. Medical educators have devised a number of educational strategies to achieve curriculum reforms based on SPICES model. For instance, to encourage problem based and active learning, tools and assessment methods were developed that demonstrate what trainees might do when faced with real clinical problems (Miller 1990). The use of clinical vignettes, case scenarios based on realistic clinical details, became important learning tools. They have been used as problem solving exercises for doctors to practice diagnosing and treating specific medical conditions (Peabody, Luck, Glassman, Jain, Hansen, Spell, and Lee 2004).

Interactive learning in small groups and web based learning has become widely used as a supplement to traditional teaching (Ruiz, Mintzer, and Leipzig 2006). Web based learning, also known as e-learning or online learning has been used for a variety of purposes such as to deliver tutorials, small group discussions, and informal assessments (Cook 2007). Clinical vignettes have been used to develop elaborate computer-based simulations of patient encounters, where students can practise taking histories, ordering and interpreting the results of diagnostic tests, and administering treatment (Cook 2007). These approaches may encourage students to understand material and take more responsibility for their own learning. Problem solving exercises and/or web based learning approaches encourage students to engage with tasks and material in meaningful ways, and the role of the educator changes from someone who teaches facts to a facilitator of learning (Ruiz et al. 2006; Barrow et al. 2010). To achieve a more integrated curriculum, different subject areas became clustered around common themes. Exposure to the clinical context was introduced in the first years to lessen the distinction between a pre-clinical and clinical phase. These integration strategies were introduced so that students could understand the commonalities between different subject areas, and how their theoretical learning relates to their professional practice in the clinical context (Barrow et al. 2010; McKimm 2010).

1.8.2. Evaluating training programmes

Evaluating the effectiveness of an innovative educational programme is an integral part of the process. The information that an evaluation provides is useful to decide whether to continue or discontinue the programme and/or identify where improvements could be made (Kirkpatrick 1998). Medical educators must demonstrate the effectiveness of their intervention using evaluation methods that are known to and valued by the medical education community (Anderson & Harris 2003). A number of frameworks have been developed to guide the design and evaluation phases of educational programmes in medical education. Some examples include Kirkpatrick's model to evaluate training programmes (Kirkpatrick 1998) and Miller's pyramid of clinical assessment (Miller 1990). Kirkpatrick recommended that the evaluation of any educational programme should be approached in a systematic manner. He proposed four broad outcomes that should be measured. These outcomes were arranged in a hierarchical sequence that requires increasing effort to evaluate. At the minimum level, an educational programme should at least assess participants' reaction to the training, followed by their learning. This could be change in knowledge or attitudes. At the higher, complex levels, training should be assessed in terms of transfer of learning to the desired context and finally long term changes at the workplace (Kirkpatrick 1998).

Miller has described four stages of learning that doctors should demonstrate in clinical assessment. At the lowest level a doctor is expected to have gained knowledge, followed by evidence they know how to use that knowledge, at least hypothetically. These refer to the cognitive aspects of learning and can be measured using traditional assessment tools such as written and oral tests. Assessing the behavioural aspects of learning becomes more difficult. This includes assessing the ability to demonstrate correct use of knowledge during practical examinations and then applying it appropriately in clinical practice (Miller 1990). This measures the ability to perform and can be achieved through observations and work place based assessments.

1.9. Assumptions underlying thesis

The preceding reviews introduce the theories, evidence and perspectives about enhancing doctors' clinical judgements and decision making skills. This thesis is based on the following perspectives: (i) doctors, like most people, are unaware of how they make judgements and decisions; (ii) they can and should be helped with formal training to make better clinical judgements and decisions; (iii) given the nature of the clinical context the emphasis of training should be on achieving well-reasoned judgements and decisions rather than desirable medical outcomes; (iv) the decision sciences offer a useful theoretical basis for intervening in doctors' clinical reasoning; and (v) interventions that aim to enhance doctors' decision literacy should be evaluated.

1.10. Thesis aims and objectives

Over thirty years has passed since research began on how to formally train doctors in clinical reasoning. The above reviews show that some training initiatives have been developed to help doctors make better clinical judgements and decisions. However it is unclear whether this type of training currently exists in medical curricula and how effective it is. Little attention has been paid to developing training courses that improve the reasoning processes doctors employ to make judgements and decisions. The aims of this thesis are: (i) to develop an intervention informed by decision theory to facilitate doctors' reasoning about clinical judgements and decisions, and (ii) to evaluate the feasibility of integrating the intervention within medical education. The objectives of this thesis include:

- identifying the evidence of interventions designed to enhance doctors' clinical reasoning;
- evaluating the effectiveness of these interventions;

- describing how interventions have been incorporated into medical curricula;
- exploring doctors' views and experiences about making clinical judgements and decisions effectively;
- making recommendations for the design and implementation of future clinical reasoning training for doctors.

The chapters within this thesis address these aims and objectives. Chapter 2 describes a systematic review integrating the evidence designed to enhance doctors' clinical reasoning. Chapter 3 describes a questionnaire study that surveyed interventions that had not been evaluated or published as research. Chapter 4 describes an interview study to explore the views and experiences of doctors in making clinical judgements and decisions. Chapter 5 describes a quasi-experimental study to evaluate the feasibility of an online decision literacy intervention.

2.

Education and interventions to facilitate doctors' clinical reasoning: a systematic review

This chapter describes a systematic review study that synthesised the evidence of interventions to enhance doctors' clinical reasoning. The chapter discusses the background, aims and objectives of the review. It then describes the details of the methods used. The results of the review are then discussed by narrative to answer the research questions. The strengths and limitations are then highlighted followed by recommendations for future work.

2.1. Background

There is increasing emphasis on the need for doctors to be taught how to be proficient in making good clinical judgements and decisions (GMC 2009). This area of training has been assumed to be acquired implicitly through the hidden curriculum, usually by opportunistic observation of reasoning demonstrated by expert doctors (Howe et al. 1984; Chessare & Lieu 1998). Some have argued that the systematic principles of human judgment and decision making should be formally taught in medical education in order to improve doctors' insight into their own clinical reasoning (Elstein et al. 1978; Elstein 1981; Elstein et al. 1982; Round 1999; Croskerry, Wears, and Binder 2000). Others have developed interventions designed to enhance specific skills such as diagnostic accuracy and treatment prescribing (Wolf et al. 1988; Hassan, Abdulla, Bakathir, Al-Amoodi, Aklan, and de Vries 2000; Akici et al. 2003).

Several reviews have integrated the evidence for the effectiveness of interventions designed to improve specific skills such as diagnosis.

accuracy and treatment prescribing. In brief, Hunt, Haynes, Hanna, and Smith (1998) integrated findings from randomised controlled trials assessing the effects of computer-based clinical decision support systems (CDSSs) on doctor performance and patient outcomes. Two-thirds of the sixty-six studies reported that computer-based clinical decision support systems improved doctors' performance for drug dosing and preventive care but the evidence was less convincing for diagnosis. Kawamoto, Houlihan, Balas, and Lobach (2005) evaluated randomised controlled trials (RCT) of decision support systems with an aim of identifying features critical for improving clinical practice. Of the seventy studies, 68% of computer-based clinical decision support systems significantly improved clinical practice. Four features were identified as active ingredients in facilitating clinical practice: automatic decision support; provision of recommendations rather than assessments; decision support at time and location of decisions making; computerised systems. Finally, Le Grand, Hogerzeil, and Haijer-Ruskamp (1999) carried out a structured review integrating findings from interventions to improve appropriate drug prescription in developing countries. Of the fifty studies reviewed, only six used an RCT design. The majority of interventions were aimed at increasing prescribing rates in public health settings rather than appropriate prescribing. Interventions such as essential drug lists and standard treatment guidelines were in wide use without being evaluated.

To date, there is no systematic review that integrates evidence of interventions aimed at improving doctors' decision literacy, such as increasing awareness of the way doctors make judgements and decisions and the factors influencing their choices. Further, it is unclear whether or not this insight into the science behind human decision making impacts on their clinical judgments, choices and practice. The review introduces the term decision literacy to describe an explicit awareness of, and ability to critically evaluate, one's own reasoning processes. This review revisits studies designed to improve both clinical judgements and decisions.

2.2. Aims and objectives

The aim of this systematic review was to integrate the evidence of strategies and interventions to facilitate doctors' clinical reasoning. The objectives were to:

- describe the interventions designed to facilitate clinical reasoning in undergraduate and postgraduate medical education;
- describe the ranges of measures used to assess clinical reasoning;
- evaluate the effectiveness of the interventions;
- identify the component parts of interventions that facilitate clinical reasoning;
- describe how these interventions have been incorporated into medical curricula;
- make recommendations for the design of future clinical reasoning interventions.

2.3. Methods

2.3.1. Design

The study was a cross-sectional survey of primary empirical research examining the interventions to facilitate doctors' clinical reasoning employing a systematic review method. There are several recognised advantages of systematic reviews over traditional non-systematic reviews (Chalmers & Altman 1995). The systematic review is an efficient method of summarising large quantities of information to make knowledge more accessible to others (Chalmers & Altman 1995; Torgerson 2003; Petticrew & Roberts 2006). They differ from traditional reviews in that they use scientific and transparent methods to identify and evaluate evidence. For example, a systematic review aims to identify all of the evidence that addresses a particular research question whereas the non-systematic review usually identifies only a subset of evidence (Torgerson 2003). The inclusion and exclusion of studies is based on a pre-defined criterion that makes the reasons for data extraction explicit (Chalmers & Altman 1995; Torgerson 2003). A coding frame is developed and used as a guide to extract the same information from each included article. This standardises the data extraction process so the researcher can compare the details and quality of the evidence (Chalmers & Altman 1995; Torgerson 2003).

These rigorous and explicit methods reduce the chances of bias influencing the results and their interpretation. The effect of the researcher's subjective opinions are minimised at all stages of the review. Their assumptions and decisions can be scrutinised by others and methods can be replicated (Torgerson 2003; Petticrew & Roberts 2006). Traditional reviews have been criticised as a haphazard overview of a convenience sample of the evidence (Torgerson 2003). As methods are not systematic, it cannot identify all of the relevant evidence that should potentially be included (Torgerson 2003). Articles are not selected on the basis of clear inclusion/exclusion criteria, this means the selection process is driven by subjectivity of the researcher (Chalmers & Altman 1995; Petticrew & Roberts 2006). Furthermore, the quality of each study is not assessed rigorously and formally as it is in a systematic review. It is unclear what interventions have been designed and which of them are effective in enhancing doctors' clinical reasoning. A thorough and systematic review of the evidence provides answers to these questions and allows suggestions to be made for future work.

2.3.2. Study selection criteria

The inclusion and exclusion criteria were revised during the course of the review in order to answer the research questions. In the initial stages the inclusion criteria were: training programmes to enhance either doctors' clinical judgements or decisions, medically trained participants, an experimental study design and publications in the English language. Excluded articles were: shared or patient decision making, participants not medically trained, case studies or those that report doctors' experiences of decision making, unpublished or non-English language research. The rationale for these exclusion criteria were as follows: the focus of the thesis is enhancing the types of judgements and decisions doctors specifically make e.g. diagnosis, prescribing and referrals. Shared or patient decision making was outside the remit of the thesis as were the clinical decisions of non-medically trained health professionals. Studies were excluded that featured shared or patients' decisions as well as participants that were not medical students or doctors. These were not relevant to understanding how to effectively enhance the clinical reasoning of doctors. Studies with a non-experimental design were excluded i.e. those without a comparative basis, qualitative and case studies. The effect of interventions could not be established without having been tested through an experiment. Studies not published or written in non-English languages were excluded due to time and financial constraints of the PhD programme.

As the review progressed three types of article were retrieved for which their inclusion or exclusion was unclear. To clarify what should happen to these studies, the inclusion and exclusion criteria were refined. The first type of article included interventions that did not require participants to engage in much thinking. For example some computer technologies and guidelines would produce the optimal decision for the participant. These articles were excluded as they did not require participants to think and make decisions for themselves. Secondly, there were many studies of medical guidelines and decision support systems. These were initially included but later excluded as their main objectives were to evaluate effectiveness of their implementation in clinical practice. Thirdly, there were studies of interventions to enhance interpretation of clinical data. In all of these types of studies, confusion arose because doctors' decision making did feature implicitly. However they were excluded on the basis that the interventions were not primarily aimed at enhancing doctors' judgements or decisions.

The final inclusion criteria were; training programmes to enhance doctors' clinical judgements or decisions, medically trained participants, experimental study design and publications in English language. The revised exclusion criteria were shared or patient decision making, participants not being medically trained, studies without an experimental design, unpublished or non-English language research, studies not requiring participants to make decisions themselves, implementation studies and studies not assessing decision making related variables. These final selection criteria reflected a more precise evidence base of training programmes to enhance the clinical reasoning of doctors.

2.3.3. Search strategy

The purpose of a search strategy is to limit the number of articles retrieved to those that are most relevant to the review (Petticrew & Roberts 2006). For this to be achieved, a balance between a sensitive (broad) and specific search was required. A preliminary search was conducted with key words in the OVID database to determine types of articles retrieved and terminology used. A few key articles were retrieved but this preliminary search produced an unmanageable number of hits, many of which were irrelevant.

To make the search more specific it required a more exhaustive list of key terms in order to retrieve relevant articles. Assistance was sought from the faculty librarian to develop a more effective strategy. The main research question was divided into the following topic headings: intervention, doctors and clinical decision making. However in order to maintain enough breadth in the search a list of synonyms for each topic heading was developed. Examples of these synonyms were training and teaching, clinicians and physicians, clinical judgement and reasoning. The purpose of including as many synonyms as possible for each topic heading was to prevent the search strategy being too specific. Bias can be introduced when the search strategy is highly specific as relevant articles can be omitted (Petticrew & Roberts 2006). Synonyms were particularly important in this review as decision making terminology is not used consistently in the medical literature. To check whether the search strategy worked effectively, the data sets retrieved from each database were screened. Relevant articles that were initially missed plus a recurrence of key articles were consistently retrieved. This suggested that the revised search strategy had achieved a balance between sensitivity and specificity. The search strategy for each database is summarised in Appendix 8.1.

2.3.4. Sources searched

The following electronic databases were searched from their start dates until May 2010: Cochrane Library, Centre for Reviews and Dissemination (CRD), Medline, Embase, PsycINFO, Cambridge Scientific Abstracts and Web of Science. The rationale for selecting these databases is discussed further below. The Cochrane Library and CRD contain the details of existing systematic reviews in the health services. These were useful resources to establish whether the same or similar systematic reviews had already been conducted. Medline and Embase were selected as they are prominent resources of medical education research. PsychINFO was selected as it consists of research articles in the Behavioural Sciences and other disciplines related to Medicine. The databases selected from Cambridge Scientific Abstracts were Applied Social Sciences Index and Abstracts (ASSIA) and Educational Research Information Centre (ERIC). ASSIA provided abstracts from social sciences journals and ERIC from education research and practice.

For a comprehensive search, it has been recommended that hand searches of selected resources should also be conducted to identify articles not found on electronic databases (Torgerson 2003; Petticrew & Roberts 2006). The journals *Quality and Safety in Healthcare*, *Medical Education* and *Medical Decision Making* were manually searched. These three journals were chosen as they publish research articles in health services, medical education and decision making theory, respectively. Research from these disciplines was considered most relevant to answering the review's research questions. The journals *Quality and Safety in Healthcare* and *Medical Decision Making* were searched from first issues in 1992 and 1981 respectively until May 2010. *Medical Education* predates both of the above journals, but the manual search began from 1981 until May 2010. The year 1981 was chosen for the noticeable increase in research on principles of clinical decision making. This work was largely headed by Arthur Elstein who encouraged the design of educational programmes that would formally train doctors in clinical reasoning (Elstein 1981; Elstein et al. 1982; Howe et al. 1984; Elstein et al. 1985). For this reason it was more likely that the majority of interventions were designed from 1981 onwards rather than before.

The work of some academics was recurring throughout the search process. In addition to the above, searches for the work of each of the following people was conducted; Arthur Elstein, Jack Dowie, Theo De Vries, Peter Ubel, Gretchen Chapman, Angela Fargerlin. Reference lists of articles included in the review were also searched so as not to miss any relevant articles.

2.3.5. Materials

A data extraction form was developed from the literature (CRD 2001) and used to extract data systematically from each article identified as meeting the inclusion criteria (Appendix 8.2). The following types of data were extracted included the following; bibliographic details, aim of study, theoretical context, study design characteristics intervention details, task details, variables measured, results and conclusions. In addition, a judgment was made on the scientific rigour of the study, i.e. quality assessment. As most of the studies in this area employ a cross-sectional and/or before-and-after and/or cohort study design, the traditional hierarchy of evidence quality assessment was not sensitive enough to discriminate the scientific rigour of the empirical research identified by this review. A set of criteria for assessing the quality of the studies was developed using guidelines from the EPPI centre (EPPI 2007). This included: the use of theory, scientific rigour, coherence of the study, and generalisable results. To differentiate between high and low quality studies, each were rated according to the above criteria and assigned a total quality score out of 30.

2.3.6. Procedure

Initially the Centre for Reviews and Dissemination databases and the Cochrane Library were searched to identify whether the same or similar systematic reviews existed. This was followed by a systematic search of the databases mentioned above. The study selection criteria were applied to potentially relevant abstracts. Full articles were retrieved when abstracts met the inclusion criteria or were in need of further analysis to make a decision. Articles not available online or from the University were ordered from the library services. Articles were categorised into three groups: include, exclude and borderline. Samples of articles to be included and excluded were reviewed by supervisors to check the suitability of the study selection criteria. All articles in the borderline category were discussed and reviewed by the supervisors in order to decide whether they met the inclusion criteria or not. During this process the selection criteria were refined to discriminate clearly between articles that should and should not be included in the review.

The data extraction form was developed and piloted on a sample of included articles to assess its suitability. After a discussion with supervisors it was simplified to reflect the wide range of study designs in medical education. The data extraction form became more general with less tick boxes and space to record study details at length. This revised data extraction form was applied systematically to each article that met the inclusion criteria of the review.

2.3.7. Analysis

The studies' characteristics were summarised in frequency tables. Statistical integration was not possible as there was little consistency in the aspects of decision making being assessed and the types of designs and measures used to evaluate effectiveness. The synthesis of findings is integrated by narrative structured in a way to address the research questions.

2.4. Results

The search identified 104,746 abstracts; 210 articles were evaluated in accordance with the review criteria; 64 were included in the final review. Figure 2 summarises the selection process of articles.

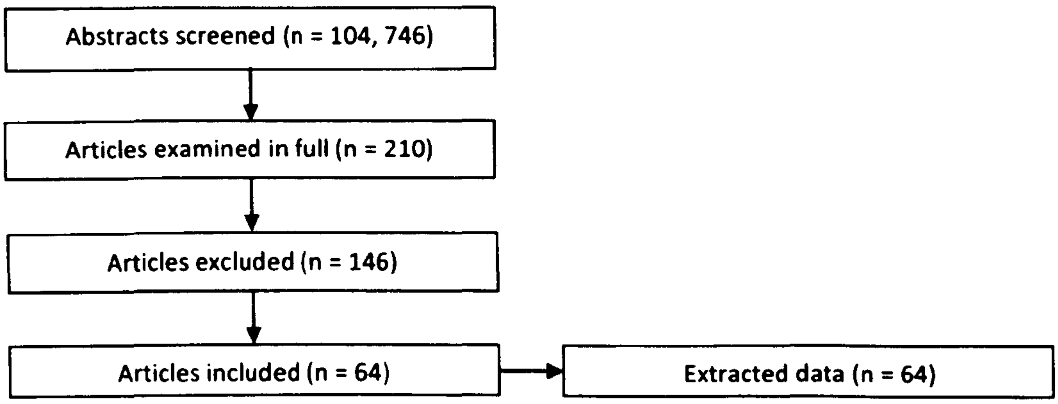


Figure 2: The selection process of articles.

Table 3 shows the study numbers and citations for all 64 articles that were included in the review.

SN	Author	SN	Author
1	Round 1999	33	Byrne et al. 2002
2	Lee, Joynt, Ho, Gin, and Hazlett 2007	34	de Dombal et al. 1991
3	Hassan et al. 2000	35	Vollebregt, Metz, de Haan, Richir, Hugtenburg, and de Vries 2005
4	Wigton, Poses, Collins, and Cebul 1990	36	DaRosa, Rogers, Williams, Hauge, Sherman, Murayama, Nagle, and Dunnington 2008
5	Friedman, Elstein, Wolf, Murphy, Franz, Heckerling, Fine, Miller, and Abraham 1999	37	Moreira et al. 2008
6	Akici, Kalaca, Goren, Akkan, Karaalp, Demir, Ugurlu, and Oktay 2004	38	Akici, Goren, Aypak, Terzioglu, and Oktay 2005
7	DeVries, Henning, Hogerzeil, Bapna, Bero, Kafle, Mabadeje, Santosa, and Smith 1995	39	Watson, Clements, Yudkin, Rose, Bukach, Mackay, Lucassen, and Austoker 2001
8	Akici et al. 2003	40	Kopp, Stark, and Fischer 2008
9	Lagerlov et al. 2000	41	Beck & Bergman 1986
10	Newton-Syms, Dawson, Cooke, Feely, Booth, Jerwood, and Calvert 1992	42	de Vries, Daniels, Mulder, Groot, Wewerinke, Barnes, Bakathir, Hassan, Van Bortel, Kriska, Santoso, Sanz, Thomas, Ziganshina, Bezemer, Van Kan, Richir, and HogerzeiL 2008

11	Veninga, Lagerløv, Wahlström, Muskova, Denig, Berkhof, Kochen, and Haaijer-Ruskamp 1999	43	Essex & Healy 1994
12	Nilsson, Hjemdahl, Hassler, Vitols, Wallen, and Krakau 2001	44	Hoffrage & Gigerenzer 1998
13	Veninga, Denig, Zwaagstra, and Haaijer-Ruskamp 2000	45	Mayou 1978
14	Schwartz, Donnelly, Nash, Johnson, Young, and Griffen 1992	46	Beullens et al. 2006
15	Lernau 1989	47	O'Connell, Henry, and Tomlins 1999
16	Carter, Butler, Rogers, and Holloway 1993	48	Frijling, Lobo, Hulscher, Akkermans, Braspenning, Prins, van der Wouden, and Grol 2002
17	Wolf et al. 1988	49	Frijling, Lobo, Hulscher, Akkermans, van Drenth, Prins, van der Wouden, and Grol 2003
18	Noguchi, Matsui, Imura, Kiyota, and Fukui 2004	50	Thomas, Boxall, Laha, Day, and Grundy 2008
19	Grant et al. 2006	51	Lincoln, Turner, Haug, Warner, Williamson, Bouhaddou, Jessen, Sorenson, Cundick, and Grant 1991
20	Abraham et al. 2004	52	Karaalp, Akici, Kocabas, Lu, and Oktay 2003
21	Hedrick & Young 2008	53	Chopra et al. 1994
22	Warner, Woolley, and Kane 1974	54	Lundborg, Wahlstrom, Diwan, Oke, Martenson, and Tomson 1999
23	Murray, Cupples, Barberm, Dunnm, Scottm, and Hannay 1977	55	Wigton, Patil, and Hoellerich 1986
24	Rogers, Grenvik, and Willenkin 1995	56	Windish, Price, Clever, Magaziner, and Thomas 2005
25	Hershberger et al. 1995	57	Tamblyn, Huang, Perreault, Jacques James, Hanley, McLeod, and Laprise 2003
26	Kurzenhauser & Hoffrage 2002	58	Anderson, McEwan, and Hrudey 1996
27	Rogers, Swee, and Ullian 1991	59	DeVries 1993
28	Servais, LaMorte, Agarwal, Moschetti, Mallipattu, and Moulton 2006	60	Sintchenko, Coiera, Iredell, and Gilbert 2004
29	Kumta, Tsang, Hung, and Cheng 2003	61	Shekelle, Kravitz, Beart, Marger, Wang, and Lee 2000
30	Cebul, Beck, and Carroll 1984	62	Gifford, Mittman, Fink, Lanto, Lee, and Vickrey 1996
31	Margolis, Barloon S, and N. 1982	63	Davidoff, Goodspeed, and Clive 1989
32	Junghans, Feder, Timmis, Eldridge, Sekhri, Black, Shekelle, and Hemingway 2007	64	Terrell, Perkins, Dexter, Hui, Callahan, and Miller 2009

Table 3: Study numbers and citations of included articles

2.4.1. Source of studies

The majority of studies were carried out in Europe (Table 4) but one third of the studies were from North America (n = 23). Studies originated from a wide range of journals but most were either specialist medical or medical education journals (Table 5).

Country	Frequency	Study Number
North America	23	5, 14, 16, 17, 19, 21, 22, 24, 25, 27, 28, 30, 36, 41, 51, 55, 56, 57, 58, 61, 62, 63, 64
Other Europe (*e.g. Netherlands, Norway, Sweden, Slovakia, Germany, Turkey, Belgium*)	20	6, 7, 8, 9, 11, 12, 13, 26, 35, 37, 38, 40, 44, 46, 48, 49, 52, 53, 54, 59
UK	10	1, 10, 23, 32, 33, 34, 39, 43, 45, 50
East Asia (*e.g. China, Japan*)	4	2, 4, 18, 29
Other (*e.g. Yemen, Australia, India, Australia, various*)	7	3, 15, 20, 31, 42, 47, 60

Table 4: Country of study origin (N = 64).

Journal Type	Frequency	Study Number
Medical specialist	25	3, 6, 8, 10, 11, 12, 13, 14, 28, 16, 19, 20, 21, 24, 32, 33, 35, 38, 42, 45, 48, 53, 54, 59, 64
Medical Education	18	1, 2, 15, 23, 26, 27, 4, 25, 17, 31, 36, 37, 40, 41, 44, 46, 52, 55
General Medicine	16	5, 7, 9, 18, 34, 39, 43, 47, 49, 50, 56, 57, 58, 61, 62, 63
Information Technology	4	22, 29, 51, 60
Decision Making	1	30

Table 5: Journal of study origin (N = 64).

2.4.2. Study characteristics

Half of the studies used a randomised controlled trial (RCT) study design (Table 6). Study 59 appears more than once in Table 6 as more than one study was conducted with different designs. The majority of studies recruited their samples using a purposive method i.e. a particular population was chosen to suit the purpose of the research (Table 7). Most studies were conducted with either medical undergraduate students only or doctors only, few used a mixed sample of both students and doctors (Table 8). One third of the studies were conducted in a university setting. Nineteen studies did not report where their study was conducted and 17 were conducted in a health care setting (Table 9).

Design	Frequency	Study Number
Randomised Controlled Trial (RCT)	32	2, 3, 7, 8, 9, 10, 11, 12, 13, 15, 22, 26, 29, 32, 33, 35, 36, 39, 40, 41, 42, 47, 48, 49, 51, 53, 54, 56, 57, 58, 62, 64
Before/after same sample	12	5, 18, 19, 20, 24, 28, 31, 34, 43, 46, 50, 52
Non-randomised comparative (experimental vs controls)	11	1, 6, 14, 16, 17, 27, 45, 55, 59, 61, 63
Before/after different samples (comparisons within same group and/or between different groups before/after)	6	4, 21, 23, 25, 30, 38
Other (designs unknown)	3	37, 44, 60

Table 6: Type of study design (N = 64).

Sampling method	Frequency	Study Number
Purposive	51	1, 2, 8, 9, 10, 11, 12, 18, 19, 20, 21, 23, 24, 25, 26, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64
Volunteer/opportunistic	9	3, 4, 5, 6, 7, 13, 22, 27, 31
Not stated	4	14, 15, 16, 17

Table 7: Sampling recruitment method (N = 64).

Sample characteristics	Frequency	Study Number
Medical undergraduates	28	1, 2, 3, 7, 14, 15, 18, 20, 22, 23, 24, 26, 27, 28, 29, 31, 35, 38, 40, 41, 42, 45, 46, 51, 52, 55, 56, 59
Postgraduates/Doctors	28	8, 9, 10, 11, 12, 13, 16, 19, 32, 33, 34, 36, 37, 39, 44, 47, 48, 49, 50, 53, 54, 57, 58, 60, 61, 62, 63, 64
Both	8	4, 5, 6, 17, 21, 25, 30, 43

Table 8: Sample characteristics (N = 64).

Study setting	Frequency	Study Number
University	22	2, 6, 7, 15, 17, 18, 19, 20, 22, 25, 26, 28, 30, 31, 35, 37, 38, 42, 46, 52, 56, 58
Not stated	19	1, 3, 4, 8, 9, 10, 14, 16, 23, 27, 29, 32, 36, 40, 41, 43, 45, 54, 55
Healthcare setting (*e.g. hospital, primary care*)	17	11, 12, 13, 24, 30, 33, 34, 39, 44, 48, 49, 50, 51, 53, 57, 63, 64
Cognitive Psychology Laboratory	2	5, 21
Other (*e.g. mailed, web*)	4	47, 60, 61, 62

Table 9: Study setting (N = 64).

2.4.3. Theoretical framework

Most studies (two thirds) did not use any theory to inform the design of their interventions (Table 10). Theory that was incorporated into some interventions was of two main types; generic and expert decision making theories. Most theories referred to were associated with improving generic reasoning skills (Table 10) i.e. theories to explain psychological processes all individuals use when making judgements and decisions. A few interventions were informed by principles of expert theories i.e. these aim to describe the psychological processes doctors use when making specialist judgements in specific contexts, such as diagnosing a condition from a set of symptoms (Table 10). As some studies incorporated more than one theory, studies can appear more than once in Table 10. No studies explicitly discussed the theories or attributes of how to make good clinical judgements or decisions.

Theory used	Frequency	Study Number
None	44	4, 5, 6, 8, 9, 10, 11, 12, 13, 20, 21, 23, 24, 28, 29, 32, 33, 34, 35, 36, 38, 39, 41, 42, 43, 45, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64
Generic decision making theory	17	1, 2, 3, 7, 15, 16, 17, 18, 19, 22, 25, 26, 27, 30, 31, 37, 44
Bayes theorem	9	1, 17, 18, 19, 22, 26, 27, 37, 44
EUT	7	2, 3, 7, 16, 27, 30, 31
Information Processing	4	1, 17, 15, 25
Expert clinical judgement theory	6	14, 15, 18, 27, 40, 46
Hypothetico-Deductive	4	15, 18, 27, 46
Pattern Recognition	2	14, 46
Schema Theories	2	14, 40

Table 10: Theoretical framework of studies.

2.4.4. Intervention Type

Interventions were categorised as either: expert heuristic, experiential learning or decision literacy (Table 11). Two thirds of the studies consisted of expert heuristic interventions i.e. those that use techniques to facilitate the specialist judgements and decisions doctors make, such as diagnosis and prescribing decisions (Table 11). These techniques were taught to people as useful strategies that could help people improve the outcome of their judgements and decisions rather than to help people understand their own reasoning processes. Fourteen studies focused on giving participants hypothetical or actual real world experience of making decisions in the clinical context i.e. experiential learning interventions (Table 11). A few interventions aimed to increase people's awareness of their own thinking i.e. decision literacy (Table 11). Over half of the studies did not integrate their interventions into existing medical courses but were carried out as one off research studies [3, 4, 5, 9, 10, 11, 12, 13, 16, 17, 21, 23, 32, 33, 34, 39, 40, 41, 42, 43, 44, 47, 48, 49, 50, 51, 53, 54, 55, 57, 58, 60, 61, 63, 64]. Twenty nine studies did integrate their interventions into medical training [1, 2, 6, 7, 8, 14, 15, 18, 19, 20, 22, 24, 25, 26, 27, 28, 29, 30, 31, 35, 36, 37,38, 45, 46, 52, 56, 59, 62]. Mostly these were integrated into existing undergraduate medical curricula rather than postgraduate training.

Type of intervention	**Frequency**	**Study Number**
EXPERT HEURISTIC	43	
Practice recommendations	25	3, 6, 7, 8, 9, 10, 11, 12, 13, 32, 35, 38, 39, 42, 43, 47, 48, 49, 50, 52, 54, 58, 59, 61, 62,
Theory based strategy (e.g. Bayes theorem)	9	2, 18, 19, 26, 27, 31, 37, 44, 63
Computer decision support	9	4, 5, 16, 34, 40, 55, 57, 60, 64
EXPERIENTIAL LEARNING	14	
Course with teaching and clinical experience	6	14, 15, 24, 36, 45, 56
Computerised scenarios, web based tutorials	6	21, 22, 23, 28, 29, 51
Simulation with manikins	2	33, 53
DECISION LITERACY	7	1, 17, 20, 25, 30, 41, 46

Table 11: Main types of interventions (N =64)

2.4.5. Variables and elicitation method

Variables measured were of two types; outcomes or processes. The majority of studies measured variables that were outcomes i.e. direct measures of the end point such as decisions, choice, knowledge, confidence and clinical skills (Table 12). The variables measured in 15 studies were of processes i.e. indirect measures of thinking such as reasoning, judgement and attitudes (Table 12). A few studies examined a mixture of outcome and process variables and these study numbers appear more than once in Table 12. Three studies did not report the variables measured and it was not possible to determine what they were (Table 12).

Many studies used more than one type of method to elicit data but there were three main types; clinical vignettes, questionnaires and patient records. Most studies obtained data via clinical vignettes and/or questionnaires (Table 13). The majority of clinical vignettes were delivered in paper format but some were also presented with use of computers (Table 13). The vignettes consisted of clinical scenarios that required participants to record their answers to open and closed questions about the scenario. Others required participants to develop a treatment plan for the hypothetical patient which was then analysed and scored for accuracy. A few studies presented the clinical scenario with a manikin [33, 53]. Participants were asked to 'work' on the manikin as if it were a real patient. One study also used the Objective Structured Clinical Examination (OSCE) where people were trained to role play patients [52].

Questionnaires were used in twenty-eight studies (Table 13). Some of them included short clinical scenarios and participants would answer associated questions in multiple-choice format (MCQ). Questionnaires without scenarios were used to obtain attitudes towards the intervention, factual knowledge and demographic information. Three studies included an observation checklist to record participants' observed behaviour [27, 36, 56]. The majority of questionnaires were study specific i.e. designed specifically for a particular study but a few used validated questionnaires. These were the Diagnostic Thinking Inventory and Inventory of Cognitive Bias in Medicine. The Diagnostic Thinking Inventory measures two aspects of diagnostic thinking across fifty-six items: the degree of flexibility in thinking and the degree of knowledge structure in memory (Bordage, Grant and Marsden 1990). The Inventory of Cognitive Bias in Medicine consists of twenty-two clinical scenarios with a choice of multiple alternative answers. For each scenario at least one choice represents a biased prone decision and one is the statistically correct decision (Hershberger, Part, Markert, Cohen, and Finger 1994).

Variable	Frequency	Study Number
Outcomes (decisions, choice, knowledge, behaviour, confidence, skills)	50	1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 21, 23, 24, 27, 28, 31, 32, 33, 34, 35, 36, 38, 39, 40, 41, 42, 43, 45, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64
Processes (reasoning, judging, attitudes)	15	1, 4, 18, 19, 20, 25, 26, 37, 43, 44, 46, 52, 54, 56
Not stated	3	22, 29, 30

Table 12: Variables measured.

Elicitation method	Frequency	Study Number
Clinical vignettes	31	
Computer	10	4, 5, 16, 21, 22, 23, 41, 51, 55, 60
Paper	21	3, 6, 7, 8, 14, 17, 24, 26, 29, 31, 32, 35, 37, 38, 39, 42, 43, 44, 45, 59, 61, 62
Questionnaire	28	
Study specific	24	2, 6, 8, 9, 11, 13, 15, 18, 19, 20, 23, 27, 28, 30, 35, 36, 38, 39, 40, 49, 54, 56, 59, 62
Validated	4	1, 25, 46, 56
Patient records	14	
Routine prescriptions	12	8, 9, 10, 11, 12, 47, 50, 54, 57, 58, 59, 64
Additional patient information	2	48, 49
Not stated	1	34

Table 13: Types of methods used to obtain data.

Timing of data collection	Frequency	Study Number
Before and after intervention	49	2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 35, 36, 38, 39, 40, 42, 43, 45, 46, 47, 4849, 50, 52, 54, 56, 57, 58, 59, 63
After intervention only	14	1, 6, 14, 15, 21, 23, 27, 30, 31, 37, 41, 53, 62, 64
During intervention	6	33, 44, 51, 55, 60, 61
Not stated	1	34

Table 14: Timing of when data was obtained.

Patient records were used in a small group of studies to obtain data (Table 13). Most were real world prescriptions used as a method of assessing doctors' prescribing habits (Table 13). These patient records were part of routine clinical practice. Comparatively, two studies required doctors to record patient information that was not part of routine clinical practice, rather these were tasks specific to the study (Table 13). These records were known as 'encounter forms' that doctors completed after each consultation with a patient. Information such as symptoms, diagnostic hypotheses and treatment given was recorded and then analysed by the researchers.

The majority of studies administered the same data elicitation method before and after the intervention to assess its effectiveness (Table 14). Those studies that used multiple methods were administered at different time points of the study; these study numbers appear more than once in Table 14.

2.4.6. Quality of studies

The quality of each study was assessed on use of theory, scientific rigour, coherence and generalisability of results. Each study was given a score out of 30. The quality of evidence across the sample of studies was not high. The majority of studies were rated as average quality (Table 15). One quarter of studies were judged to be of quite poor or poor quality (Table 15). The main reason these studies received low quality scores was due to a lack of sufficient methodological detail. This made it difficult to understand the nature of the interventions, e.g. type of tasks participants were given and when, validity of measures used. However, it is important to recognise that poor reporting of methods does not necessarily mean that the interventions have been poorly designed (CRD 2009). For this reason, studies judged to be of low quality were included in the review's analysis. It does mean that some of the conclusions from the review are partly based on poor evidence and therefore need to be accepted with caution e.g. five effective and six partly effective interventions were judged to be of poor quality.

The following methodological flaws were identified consistently across the sample of studies; poor use of theory, incomplete scientific rigour, limited generalisable results and poor coherence. Of the twenty interventions that were informed by theory, only thirteen explicitly referred to an established theory i.e. they named theories and discussed how they were being operationalised [1, 2, 14, 17, 18, 19, 22, 25, 26, 27, 37, 40, 44]. However, none of these studies discussed or evaluated the theories they had used in the discussion section of the articles (Table 16). Seven studies referred to a theory implicitly but did not describe how it was operationalised [3, 7, 15, 16, 30, 31, 46].

Quality score	Frequency	Study Number
Average (20-24)	43	3, 5, 6, 8, 10, 11, 12, 14, 16, 19, 20, 21, 23, 24, 25, 26, 28, 30, 31, 32, 35, 36, 37, 39, 40, 42, 44, 46, 47, 48, 49, 51, 52, 53, 54, 56, 58, 59, 60, 61, 62, 63, 64
Quite poor (15-19)	13	4, 9, 18, 27, 29, 33, 34, 38, 41, 43, 50, 55, 57
Good (25-30)	5	1, 2, 7, 13, 17
Poor (10-14)	3	15, 22, 45

Table 15: Quality of study scores (N = 64).

Criteria of quality assessment	Frequency	Study Number
Poor use of theory	56	3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 32, 33, 34, 35, 36, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64
Incomplete scientific rigour	47	2, 3, 4, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26.27, 28, 29, 34, 37, 38, 39, 41, 42, 43, 44, 45, 47, 48, 50, 52, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63
Lack of generalisable findings	32	3, 8, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 25, 27, 28, 29, 30, 31, 37, 40, 41, 44, 45, 46, 50, 51, 53, 55, 58, 60, 61, 63
Poor coherence	25	4, 9, 12, 15, 16, 22, 24, 26, 27, 28, 29, 30, 31, 33, 34, 37, 40, 43, 45, 48, 51, 55, 56, 57, 59

Table 16: Criteria for quality assessment of studies.

Generally, the scientific rigour of studies was incomplete (Table 16). In some studies the appropriateness of the sample was difficult to determine and the majority of studies did not report the validity and reliability of their measurements. Those that did commented on either

validity [1, 2, 13, 40] or reliability [33, 40, 46, 49] but not both. A few studies reported at least one of their measures was both valid and reliable [13, 14, 25, 30, 36, 56] but only seven reported evidence of validity or reliability with corresponding coefficients [30, 36, 40, 46, 49, 51, 56]. Only 9 studies that used clinical vignettes reported they were reviewed by a team of experts of which content validity can be assumed [5, 23, 29, 32, 35, 59, 60, 61, 62]. Fewer studies reported that their clinical vignettes had been piloted before use [61, 62].
The generalisation of findings in half of the studies was difficult to ascertain as it was not clear if the sample was representative and/or limitations of the study were not acknowledged (Table 16). The coherence or logical consistency of 25 studies was also difficult to determine as methods and results were poorly written (Table 16).

2.4.7. Effect of intervention

Thirty one studies reported significantly improved results in the experimental group compared to the control group on all outcome variables measured (Table 17). Twenty-six studies were partly effective (Table 17) when the intervention was associated with (i) significantly better results on some but not all outcomes, (ii) improved results but no significant differences and (iii) improved results but statistical evidence not reported. Seven studies reported their intervention had no effect on outcomes measured (Table 17).

Effect of intervention	Frequency	Study Number
Effective	31	1, 3, 7, 8, 9, 10, 16, 17, 19, 20, 21, 23, 24, 25, 28, 29, 30, 32, 36, 38, 40, 43, 46, 50, 52, 53, 58, 60, 61, 63, 64
Partially effective	26	2, 4, 5, 6, 11, 12, 13, 14, 26, 31, 34, 35, 39, 41, 42, 44, 45, 48, 49, 51, 54, 55, 56, 57, 59, 62
Not effective	7	15, 18, 22, 27, 33, 37, 47

Table 17: Effect of intervention (N = 64).

2.4.8. Effectiveness and other variables

The following section examines whether relationships exist between effectiveness of interventions and other variables. Effectiveness is analysed in relation to: type of intervention, integration of interventions, participants' stage of medical career, presence of feedback and quality of study.

Type of intervention	Effective	Partly effective	Not effective	Total
EXPERT HEURISTIC				
Practice recommendations	3, 7, 8, 9, 10, 32, 38, 43, 50, 52, 58, 61, 64	6, 11, 12, 13, 35, 39, 42, 48, 49, 54, 59, 62	47	**26**
Theory driven strategies (decision analysis, Bayes theorem)	19, 63	2, 26, 31, 44	18, 27, 37	**9**
Computer decision support	16, 40, 60	4, 5, 34, 55, 57		**8**
EXPERIENTIAL LEARNING				
Computerised tutorials	21, 23, 28, 29	51	22	**6**
Course with teaching and clinical experience	24, 36	14, 45, 56	15	**6**
Simulation with manikins	53		33	**2**
Decision Literacy	1, 17, 20, 25, 30, 46	41		**7**
Total	**31**	**26**	**7**	

Table 18: Effectiveness and intervention type.

Interventions that were most associated with effectiveness were those that taught practice recommendations to participants, computerised tutorials of decision making problems and teaching decision literacy (Table 18). The majority of these interventions were effective or partly effective with only a couple not being effective [22, 47]. The findings from the decision literacy interventions were particularly encouraging. Six out of seven were reported to be effective (Table 18).

Interventions that had been integrated into medical training were no more effective than those that had not been integrated. The majority of the one off research studies were effective or partly effective; only two were not effective (Table 19). Similarly, most interventions integrated into medical curricula were effective or partly effective but five were not effective (Table 19). This suggests that more of the one off research studies were effective compared to those integrated into medical training.

Integration of study	Effective	Partly	Not effective	Total
One off study	3, 9, 10, 16, 17, 21, 23, 32, 40, 43, 50, 53, 58, 60, 61, 63, 64	4, 5, 11, 12, 13, 34, 39, 41, 42, 44, 48, 49, 51, 54, 55, 57	33, 47	**35**
Integrated into medical education	1, 7, 8, 19, 20, 24, 25, 28, 29, 30, 36, 38, 46, 52	2, 6, 14, 26, 31, 35, 45, 56, 59, 62	15, 18, 22, 27, 37	**29**
Total	**31**	**26**	**7**	

Table 19: Effectiveness and integration of interventions.

Effect of intervention	Medical undergraduates yrs 1 & 2	Medical undergraduates yrs 3+	Doctors	Total
Effective	20, 36	1, 3, 21, 23, 24, 25, 28, 29, 30, 38, 43, 46, 52,	8, 9, 10, 16, 19, 32, 50, 53, 58, 60, 61, 63, 64	**28**
Partly effective	4, 26, 55, 56	2, 5, 6, 31, 35, 51, 59	11, 12, 13, 34, 39, 44, 48, 49, 54, 57, 62	**22**
Not effective	22	15, 18, 27	33, 37, 47	**7**
Total	**7**	**22**	**26**	

Table 20: Effectiveness and stage of participants' medical career.

Not all studies reported details of participants' stage of medical career. Those that did (n = 55) showed that most effective and partly effective interventions were conducted with medical undergraduates beyond their second year and/or doctors of various levels of experience (Table 20). However this pattern may be misleading as not many studies were conducted with medical undergraduates below their third year (Table 20). Furthermore, the few that did include first and second year medical students were effective or partly effective; only one was not an effective intervention (Table 20).

Two thirds of studies offered feedback to participants on their learning and/or performance (Table 21). However these studies were no more associated with effective interventions than those that did not offer feedback; the majority was effective or partly effective irrespective of whether feedback was given or not. Furthermore, most non-effective interventions did offer feedback to participants (Table 21). This suggests that the type, rather than the presence of feedback during the study may impact on effectiveness of interventions.

Effect of intervention	Feedback given	Feedback not given	Total
Effective	1, 3, 7, 8, 9, 20, 28, 29, 30, 36, 38, 40, 46, 50, 52, 53, 63	10, 16, 17, 19, 21, 23, 24, 25, 32, 43, 58, 60, 61, 64	31
Partly effective	4, 5, 6, 11, 12, 13, 14, 31, 35, 42, 48, 49, 51, 54, 55, 56, 57, 62	2, 26, 34, 39, 41, 44, 45, 59	26
Not effective	15, 22, 27, 33, 37, 47	18	7
Total	41	23	64

Table 21: Effectiveness and provision of feedback.

Study quality was similar irrespective of whether the intervention was effective or partly effective. Most of these studies were judged to be of average quality but a few were quite poor (Table 22). This implies that effective interventions were not associated with better study quality than partly or non-effective interventions. However, study quality of non-effective interventions was mostly judged to be quite poor or poor (Table 22). Few were average and none were judged to be of good quality. Generally the non-effective interventions were associated with more methodological flaws than other studies i.e. low quality scores.

Effect of intervention	Good	Average	Quite poor	Poor	Total
Effective	1, 7, 17	3, 8, 10, 16, 19, 20, 21, 23, 24, 25, 28, 30, 32, 36, 40, 46, 52, 53, 58, 60, 61, 63, 64	9, 29, 38, 43, 50	0	31
Partly effective	2, 13	5, 6, 11, 12, 14, 26, 31, 35, 39, 42, 44, 48, 49, 51, 54, 56, 59, 62	4, 34, 41, 55, 57	45	26
Not effective		37, 47	18, 27, 33	15, 22	7

Table 22: Effectiveness and study quality

2.5. Discussion

This review systematically evaluated 64 primary empirical studies of interventions to enhance doctors' clinical reasoning. It summarises the types of training interventions and their effectiveness, the range of judgement and decision making measures, integration into medical curricula and recommendations for the design of future interventions. The review indicates that there is limited quality research evidence on how to acquire good clinical reasoning skills. The following discussion sections are based on studies that were judged to be of average and/or good quality. Studies that received the lowest quality scores are not referred to as they offered little useful and reliable information about how to effectively enhance doctors' clinical reasoning skills.

There is little formal training for doctors' in clinical reasoning in the undergraduate curricula. This is disappointing considering it was recognised more than thirty years ago that it can and should be included in medical education (Elstein et al. 1978). However, the review found that the integration of an intervention into medical curricula was not associated with effectiveness. This suggests that simply including clinical reasoning training in medical education is not enough. Rather, the timing and place need to be carefully considered to produce optimum benefit (Chessare & Lieu 1998).

It was disappointing that only a few authors explained the rationale to how they integrated their intervention into medical education and/or suggestions for how future interventions should be integrated. One suggestion was that learning needs to be reinforced at various stages throughout medical training rather than incorporate clinical reasoning training at just one point in the curriculum (Rogers et al. 1991; DeVries 1993; DeVries et al. 1995). A specific recommendation was to encourage development of knowledge and its application simultaneously, rather than a traditional sequential format of knowledge gain and the learning of its application (DeVries et al. 1995). Preclinical students effectively learned therapeutic problem solving without needing to master relevant pharmacology knowledge first before gaining clinical experience in making prescribing decisions (Abraham et al. 2004; Vollebregt et al. 2005). These findings imply that aspects of medical education traditionally delivered in the final years of curriculum may be introduced earlier (Vollebregt et al. 2005).

Most interventions trained doctors to employ specific expert strategies that would improve diagnostic or treatment decisions (expert heuristics). Most interventions measured effectiveness by an outcome such as improvements in diagnostic accuracy, a more cost-effective treatment or a component of knowledge. There was little emphasis on improving the reasoning processes that doctors use to reach judgements and decisions. However, the few that delivered courses to enhance doctors' awareness of their own decision making processes (decision literacy), were effective. Doctors' clinical judgements and decisions improved after training in critical thinking and/or an awareness of how errors occur in the reasoning process that can lead to biased decisions.

Training doctors to become more decision literate involves understanding about the generic aspects of human judgement and decision making that is applicable to all types of doctors and clinical decisions. Learning should be accessible over time and in different contexts (Metcalfe & Shimamura 1994) and the present review demonstrates that the majority of interventions have been designed to facilitate specific decisions or parts of the process to reach a judgement or decision. These types of interventions limit the transfer of learning to other contexts. Training doctors to have an awareness and ability to critique their own reasoning processes may be a way to overcome this limitation.

It was unclear as to which component parts of interventions are associated with the enhancement of clinical reasoning skills. However based on existing literature it is recommended that future interventions are designed with the following factors in mind.

- **Theoretical basis:** Interventions should be designed with reference to a theoretical basis so that underlying assumptions are explicit and possible mechanisms that contribute to the effectiveness of interventions can be identified (Chen 1990; MRC 2008). It is likely that interventions will have to draw on evidence from the decision sciences on how individuals and experts make judgements and decisions and the factors influencing their choices.

- **Appropriate assessments:** Medical education should employ assessment methods targeting not only knowledge components such as recall tests but methods that demonstrate what trainees will do when faced with real clinical problems (Miller 1990). There should also be some attempt at assessing the long-term impact of training and transfer of learning to clinical practice.
- **Informative feedback:** Feedback is an important aspect of successful learning and delivering informative feedback is a skill itself (King 1999; Race 2005). The review found there was an association between type of feedback given to participants and effectiveness of intervention. Good feedback was constructive, individualised to each participant and was encouraging of reflection and discussion (Chopra et al. 1994; DeVries et al. 1995; Akici et al. 2003; DaRosa et al. 2008). On the other hand, description of performance rather than explanation and no opportunity for participants to discuss their performance with others was likely to be less useful (Rogers et al. 1991; O'Connell et al. 1999; Byrne et al. 2002).
- **Integrate into medical education:** There is little value in designing a clinical reasoning intervention without at least considering where it could be placed in medical education. Issues such as where to place clinical reasoning training and who could teach it require careful consideration. It is important to have people delivering a new course who will champion it and imaginative ways to deliver them should be sought to encourage staff and students to take it seriously (Cuff & Vanselow 2004).

2.6. Critique of Study

The strengths of this review are the use of a rigorous method to integrate systematically the range of disparate research in this area. The broad set of search terms used ensured that a range of different types of clinical reasoning interventions were evaluated. Whilst an effort was made to develop a sophisticated search strategy that reflected the interchangeable use of decision making terminology, there is a chance that some key words may have been missed during the search. Due to time and financial constraints the review was limited to studies that were published and reported in the English language. Therefore studies that have not been published and/or were reported in other languages were not reviewed.

2.7. Summary

Few studies have evaluated the effectiveness of training to enhance doctors' clinical reasoning. Overall, the quality of evidence reviewed was not high due to poor written presentation of methods and lack of scientific rigour. Although decision literacy interventions were few in number, the findings from this type of training were encouraging. Future research should consider improving doctors' decision literacy as a potentially successful way of enhancing the quality of their clinical judgements and decisions. It is not clear which components are associated with an effective clinical reasoning intervention. However the design of future interventions should consider the following guidance based on existing literature: explicit use of established theories of judgement and decision making to help contextualise and enable doctors to reason more explicitly about their choices; use of appropriate assessment methods that reflect learning objectives; incorporation of informative feedback into the learning experience and integration or at least consideration of where interventions can be integrated formally into medical education. A consensus should also be reached about the usage of the terminology used in the clinical reasoning literature. Currently, elements of training doctors in clinical reasoning appear under a variety of curriculum topics like communication skills, evidence based medicine, and critical appraisal. These elements are not explicitly identified or labelled as topics aimed to improve clinical judgements or decisions. It advised that conclusions are taken with caution as they are partly based on a selection of poor evidence.

3.

Education and interventions to facilitate doctors' clinical reasoning: a questionnaire study

This chapter describes a questionnaire study to survey the training courses designed to improve doctors' clinical reasoning. Chapter 2 described a systematic review that was conducted to integrate the evidence of training interventions designed to facilitate doctors' clinical reasoning. The review focused only on interventions that were evaluated and published as studies in peer reviewed journals. From discussions with supervisors and colleagues who had an interest in improving doctors' clinical judgements and decisions, it was recognised that other courses may have been developed that had not been evaluated and/or which had not been published as research. It would have been inappropriate to summarise the evidence based published work only. As unpublished work was not included in the systematic review, this questionnaire study was conducted to extend previous findings reported in Chapter 2.

3.1. Background

Research into the formal training of doctors to improve their clinical reasoning has been underway for at least thirty years. This work has previously been surveyed to summarise details of training courses and document trends (Elstein 1981; Elstein et al. 1985). In 1980, Elstein distributed questionnaires to 100 members of the Society for Medical Decision Making (SMDM). The aims were (i) to develop a network of contacts who were delivering clinical reasoning training to doctors and (ii) to survey the preferences for a variety of possible clinical reasoning courses (Elstein 1981). The number of people found to be offering such courses was low. Of the 80 respondents, 16 reported that part of their training focused on improving clinical reasoning, whilst 28 said they offered formal courses in clinical reasoning. Respondents considered the most useful way to deliver this type of training to be through workshops, a dedicated journal and conferences. Medical students, junior doctors and medical school staff were thought to be the most important people to train.

The SMDM members were surveyed again a few years later with two further questionnaires about clinical reasoning training (Elstein et al. 1985). These questionnaires focused on obtaining information about the structure and content delivered in their courses as well as topics of interest for future training. The combined results of both questionnaires indicated an increase in the number of individuals and/or institutions offering clinical reasoning training. There was noticeable emphasis on delivering some training in decision analysis. The majority of respondents stated they spent some time teaching principles and techniques of decision analysis delivered usually via lectures, informal rounds or case conferences. There was a consensus that a good clinical reasoning course teaches trainees the techniques to apply decision theory such as how to use decision trees, Bayes calculation and 2×2 contingency tables (see Section 1.7).

3.2. Aims and objectives

Surveys of previous work on clinical reasoning training provide a knowledge base for developing future training in this area for doctors. Over thirty years has passed since a survey of doctors' clinical reasoning training was conducted. The aims of the present study were to (i) extend findings from the systematic review study in Chapter 2, by surveying courses not evaluated and/or published in peer reviewed journals and (ii) to extend the findings of Elstein and colleagues by surveying a broader specialist audience beyond members of the Society for Medical Decision Making. The objectives were to:

- describe the types of training courses in clinical reasoning for doctors in undergraduate and postgraduate medical education;
- identify whether courses have been incorporated into medical curricula;
- describe views and preferences of how clinical reasoning courses should be designed.

3.3. Method

3.3.1. Design

This study used a cross-sectional survey design with a questionnaire method to elicit data. There are several advantages of using a questionnaire-based survey (Robson 2002). They are relatively simple to conduct and provide an efficient way of collecting a large amount of standardised data (Robson 2002; Schutt 2006). It is also a feasible way to survey the attitudes, beliefs and behaviours from a wide range of geographical locations. There is opportunity for questionnaires to be self-completed anonymously by respondents. This can help to reduce the influence the researcher has on the findings of the study. For example, self completion can lessen the tendency towards socially desirable responses, a potential problem associated with interviews (Robson 2002).

However, the questionnaire method is associated with limitations such as low response rates and problems with the validity of data obtained. To encourage participation, the researcher needs to put considerable time and effort into developing a suitable questionnaire format and recruitment strategy. Ensuring that a questionnaire is widely distributed and potential volunteers are contacted with follow up reminders, are ways to increase the response rate (Coolican 1999). As the researcher has no control over how the questionnaire is completed, the validity and quality of data obtained can be limited by incorrect responses to questions and/or missing answers (Robson

2002). These problems may occur due to misunderstanding a question or not completing the questionnaire seriously. It is possible to increase the accuracy of the results by piloting the questionnaire. This helps to check that the questionnaire is worded clearly and the format is suitable to facilitate appropriate responses. In this study, a questionnaire method was considered to be a feasible way to survey an international community of medical educators about clinical decision making training.

3.3.2. Ethical approval

The study received ethical approval in 2009 from the University of Leeds Faculty of Medicine and Health Research Ethics Committee (see Appendix 8.3).

3.3.3. Sample

Medical educators that had taught doctors about an aspect about improving clinical reasoning were invited to complete the questionnaire. A purposive sampling method was used to recruit participants. In purposive sampling, the researcher selects the sample in a deliberate and non-random manner in order to fulfil the study's objectives (Robson 2002). A limitation of this method is that it may not truly represent the population of interest (Robson 2002). However it was appropriate in this study as it allowed the questionnaire to be distributed specifically amongst medical educators, who may have been involved in delivering clinical reasoning training to doctors. The questionnaire was distributed to members of the following groups; Evidence Based Shared Decision Making, Association for the Study of Medical Education, European Association of Decision Making and Society for Medical Decision Making. These groups were selected as many of the members were known to be involved in delivering medical education and/or had a specialist interest in doctors' improving clinical reasoning.

3.3.4. Materials

3.3.4.1. Cover letter
A cover letter was written to accompany the questionnaire. This explained to potential volunteers the study's purpose, how to participate and how long it was expected to take to complete the questionnaire. It also informed them about how the data would be used and that confidentiality and anonymity of participants would be maintained. The cover letter was addressed from the author's supervisor who was a member of some of the target groups (see Appendix 8.4). This was to encourage people to complete the questionnaire.

3.3.4.2. Questionnaire
One aim of this study was to extend the findings from the systematic review study in Chapter 2. To reflect this, the results from the systematic review were used to develop the focus of the questionnaire. Several draft versions of the questionnaire were developed and discussed with supervisors in order to achieve a suitable content and structure. A final version is appended as Appendix 8.5. A summary of this development phase is described below.

3.3.4.3. Construction of questions
The systematic review provided information about the following broad areas listed below. These areas were used as a basis to form related questions on the questionnaire.

- types of training interventions
- focus and content of training interventions
- theoretical framework
- methods of delivery
- effectiveness of training interventions
- timing of training interventions

In order to increase the validity of the questionnaire, as well as the response rate, questions were developed according to the following recommendations. The total number of questions was minimised so that the length of the questionnaire would be no longer than two pages and individual questions were written as succinctly as possible. This was to ensure that participants could complete the questionnaire quickly (within approximately five minutes) and with ease (Robson 2002).

An effort was made to phrase instructions and questions using clear and simple language to reduce potential misunderstandings (Coolican 1999; Robson 2002; Schutt 2006). Instructions on how to respond to each question were made explicit such as *tick all that apply* or *tick one response.* Further, each question was phrased in such a way that only one question was asked at one time and in a neutral manner. This was to maintain clarity and so as not to introduce bias (Coolican 1999; Robson 2002; Schutt 2006). For example, questions regarding the content of a course were phrased separately in the following way;

> What is the focus of your medical decision making course(s)?
>
> What theoretical perspectives do you draw on during your teaching of the course(s)?
>
> What applications do you refer to during your teaching of the course(s)?

A wide range of response options were listed to ensure data were as representative of teaching in this area as possible (Coolican 1999; Robson 2002; Schutt 2006). To ensure all responses were available the option 'other' was presented where applicable. A neutral option was available on rating scales so as not to create opinions and an opportunity was given to leave further comments at the end of the questionnaire.

3.3.4.4. Structure of questionnaire

The questionnaire was structured so that it appeared clear and simple to complete for participants (Robson 2002; Schutt 2006). In addition to a cover letter, a brief introduction was written at the top of the first page of the questionnaire. This summarised the purpose of the study, the participant's task, and information about the author. The text was formatted in such a way that the title, instructions, questions and response options were distinguishable from one another. The response options for each question were presented in a table. This allowed participants to respond using the familiar tick box format (Robson 2002). Where response options were presented on a rating scale these were structured in a matrix to standardise the layout (Schutt 2006). This was appropriate for questions 9 and 10. These questions were structured as a series of statements associated with a common question and the same response format for each statement. To give the questionnaire a logical framework, questions were ordered in a sequence. This was done by separating the questions into two main themes. The first page of the questionnaire focused on obtaining factual information about an existing clinical reasoning course. This included questions about how and where it was delivered, and the content of the course. Whereas the second page of the questionnaire focused on participants' views of what they think makes a good clinical reasoning course. This included questions around how to deliver teaching, content, timing and assessment.

3.3.5. Pilot study

The questionnaire was piloted at the thirty-first annual meeting for the Society of Medical Decision Making (SMDM). This was considered appropriate as it attracts an international audience of medical educators who have a specialist interest in doctors' clinical reasoning. The conference organisers agreed to accommodate the questionnaire by including a printed version in each of the delegates' conference packs. The questionnaire was tested for clarity, comprehension and distribution method. The pilot study did not result in any major changes to the questionnaire. Participants completed the questionnaire correctly so it was assumed that the format and instructions of the questionnaire were clear. However, the response rate was low with only five people completing the questionnaire. One possible reason for this was that it was not convenient for people to complete a questionnaire at a meeting with a demanding schedule. Thereafter the questionnaire was sent electronically to the Society of

Medical Decision Making distribution list, where more people responded. This suggested that people found it more convenient to complete an electronic rather than paper version of the questionnaire.

3.3.6. Procedure

The majority of delegates at the SMDM meeting did not respond to the questionnaire. To increase response rates the survey was e-mailed to members of the SMDM. This usually incurs a fee however a sub-group known as the Diagnostic Errors in Medicine group, agreed to pay this cost by sponsoring the study. Electronic versions of a cover letter and questionnaire were e-mailed to members of the SMDM. The following groups also agreed to accommodate the study where the materials were also distributed electronically; members of the Association for the Study of Medical Education's Education Research Group, Evidence Based Shared Decision Making and European Association of Decision Making. For each group, the materials were distributed by a member of the societies' administrative team. Potential participants were instructed to complete the questionnaire anonymously and return it via e-mail by the specified date. To increase the response rate the respective administrators sent a follow up e-mail to each group that reminded people about the study.

3.3.7. Analysis

The data were summarised in frequency tables. This was considered suitable as the aims of the study were to describe clinical reasoning courses. Statistical analysis was not performed as the study was not interested in comparisons between participants' responses or correlations between variables. For each question a frequency table was constructed that described the number of people that selected each of the associated options. This clarified the most and least common practices and opinions of the sample group.

3.4. Results

Questionnaires were completed by 40 individuals involved in teaching doctors about an aspect of clinical reasoning. The following sections summarise the frequency of responses to each questionnaire item as well as total number of responses, missing responses and multiple responses i.e. where more than one response option was selected. Results are presented according to the study's objectives.

3.4.1. Types of clinical reasoning courses

Most of the courses took place somewhere in the USA and/or Canada, with just over half of all respondents selecting this location. Roughly a third of respondents stated that their training is offered in a European country. A few respondents did not state which country their training takes place and one respondent stated their training happens in various countries.

The vast majority of training was delivered in a university and a third of courses took place at conferences. Less than a third offered training for medical professional bodies and public health facilities. Examples of types of medical professional bodies were the General Medical Council and the Norwegian Medical Association. Examples of public health facilities were hospitals and clinical and research ethics committees. Training at other organisations included a cancer helpline and resource centres for patient support. Almost half of the respondents stated that they have offered clinical reasoning training at more than one type of organisation.

Focus of course	Freq.	Theory used	Freq.	Applications referred to	Freq.
Helping doctors make better clinical judgements and decisions	29	Classical/normative	37	Risk presentation	23
Helping doctors skills in patient centred care or shared decision making	29	Information processing/heuristics and bias	28	Decision analysis	22
Helping doctors develop general awareness of clinical reasoning	23	Expert models of decision making/naturalistic	12	Bayes theorem	21

Helping doctors make better team judgements and decisions	11	Social cognition models	7	Patient decision aids	21
Informing doctors about social sciences	7	Self regulation theories	4	Utility elicitation methods	14
Other	6	None	2	Smart heuristics	9
Total responses	*40*	Other	9	Other algorithms	3
Multiple responses	*30*	*Total responses*	*40*	Other	3
		Multiple responses	*33*	*Total responses*	*39*
				Missing responses	*1*
				Multiple responses	*33*

Table 23: Content of clinical reasoning courses.

Clinical reasoning courses most commonly focused on helping doctors make better clinical judgements and decisions and/or developing skills in patient centred care or shared decision making. Just over half of all courses focused on helping doctors to develop a general awareness of their clinical reasoning (Table 23). Examples of other responses included teaching researchers about studying decision making, enhancing ethical decision making and teaching about a new model developed from neuroscientific evidence. The majority of courses were based on one theoretical perspective or more (Table 23). The commonest theories referred to were a combination of both classical decision making theories and information processing. Examples of responses that people referred to as other theories were Risk Communication, Moral Reasoning Theories, Evidence Based Medicine, Systemic and Constructivist Communication Theory, and the Ottawa Decision Support Framework (Table 23). The majority of respondents stated that they referred to more than one application in their course, the most common being risk communication, decision analysis, Bayes theorem and/or patient decision aids.

3.4.2. Integration into medical curricula

Nearly half of the respondents stated that they deliver clinical reasoning courses in more than one way (Table 24). Approximately two thirds of respondents included some teaching on clinical reasoning in medical curricula, usually delivered as a module on clinical reasoning specifically making or a module on a broader topic (Table 24). Modules on broader topics included shared decision making, communication skills, problem based learning and evidence based learning. Of those that had included the training in medical curricula, it is not known how many courses were temporary or permanent features in medical education. Nearly half of the respondents stated that training was delivered as specialist courses, most of which lasted for half a day. These were not usually part of medical curricula. They were more commonly aimed at introducing principles of decision making to other health professionals that were not doctors, usually at conferences or their workplace (Table 24). Other delivery methods of courses included short introductory or overview lectures to students and/or doctors, longer specialist courses ranging from three to five days and a fifteen week online course.

Delivery Method	Frequency
Module on medical decision making	15
Part of module on broader topic	14
Half day specialist course	12
One-two day specialist course	6
Other	9
Total responses	40
Missing responses	0
Multiple responses	18

Table 24: Delivery method of teaching clinical reasoning.

3.4.3. Views and preferences towards future courses on improving clinical reasoning

There was no topic that was considered to be the key feature to include in teaching about clinical reasoning. Most options were selected by over half of all respondents each time. Fifteen respondents thought that all, or nearly all, of the topics should be included in a clinical reasoning course (Table 25). Similarly, there was also no single method of teaching that was considered the most important way of teaching to improve clinical reasoning (Table 25). Most respondents thought that courses should be delivered using more than one method (n = 33). Nine people thought that all, or nearly all of the above methods of teaching were important to use (Table 25). Examples of other methods suggested were role plays, interactive group seminars, positive critical incidents and video recording with feedback.

Topics to include	Frequency	Method of teaching to use	Frequency
Risk perception	29	Real patient care experience	28
Probabilities	27	Paper based clinical scenarios	25
Decision aids	27	Standardised patients/actors	20
Diagnostic test accuracy	26	Erroneous clinical examples	14
Decision analysis	24	Anecdotes of doctors	14
Clinical practice guidelines	24	Computerised clinical scenarios	13
Clinical reasoning strategies	24	High fidelity simulation	7
Heuristics and bias	23	Other	11
Cost effectiveness analysis	14		
Other	10		
Total responses	*40*	*Total responses*	*39*
Multiple responses	*38*	*Multiple responses*	*33*
Missing responses	*0*	*Missing responses*	*1*

Table 25: Importance of topics and teaching methods for future courses.

Table 26 summarises how relevant people thought each of the above statements were towards clinical reasoning courses. Some people did not respond to every statement, with a range of three to eight responses missing from some questionnaires. Most people strongly agreed with integrating a course on clinical reasoning into professional training post first degree but that if it is to be included in undergraduate training then integrated within communication skills and/or during the clinical attachment phase is the most appropriate time (Table 26). All respondents showed some agreement with the inclusion of clinical reasoning training as a means of increasing patient safety, however over 50% believed it was unlikely to gain wide support in medical education (Table 26).

Statements	Not relevant	Possibly relevant	Very relevant
Integrated with social sciences courses in undergraduate training (n = 34)	32%	41%	26%
Integrated into communication skills courses in undergraduate training (n = 36)	11%	25%	64%
Integrated with clinical attachment in undergraduate training (n = 32)	6%	31%	63%
An optional specialist topic in undergraduate training (n = 32)	34%	44%	22%
Integrated into professional training post first degree (n = 33)	3%	12%	85%
An optional component in continuing professional development portfolios (n = 33)	12%	48%	39%
A work based learning activity (n = 37)	5%	50%	49%
Doctors' decision making should be assessed in the workforce (n = 35)	3%	31%	66%
Formal decision making training will increase patient safety (n = 37)	0	50%	54%
It is unlikely formal decision making training will gain wide support in medical education (n = 33)	42%	33%	24%

Table 26: Views about when to teach clinical reasoning and its impact on medical education.

3.4.4. Advice to medical educators

Eighteen respondents offered advice to other medical educators regarding how to approach the design and implementation of future clinical reasoning courses. These were summarised into the following three themes: (i) what to teach (ii) how to teach and (iii) where to place the course in medical education.

Advice on what to teach: There was a range of general and specific suggestions by seven respondents. More than one respondent stated that enhancing doctors' understanding of risk was important and one suggestion was to present risk information in different formats. A few people thought doctors need to be encouraged to use decision aids more, whereas one person stated that decision analysis and decision aids should only be introduced as a conceptual tool (i.e. little practical relevance). The importance of teaching ways to facilitate patient involvement in decision making was raised by more than one person. Two people from the same department believed that the role of ethics in decision making is not given enough importance. Their advice was to train doctors to be aware of the ethical issues that arise in practice. Three people gave advice on teaching some generic aspects of decision making. Examples were (i) focusing on understanding one's own reasoning with less emphasis on memorising facts so that processes are understood and decisions can be justified, (ii) emphasising the problem of bias inherent in human judgement and decision making.

Advice on how to teach: Nine respondents gave suggestions about how clinical reasoning courses should be taught. These were mostly broad suggestions about the importance of using clinical scenarios to allow doctors to practice making judgements and decisions. Examples were via role plays, computers and simulated patients. Five respondents highlighted the importance of increasing doctors' perceived relevance of clinical reasoning training. Two people noted that motivation towards participating in such courses is low and that medical educators should invest time in raising awareness about the importance of good clinical reasoning as a lifelong habit. It was stressed that imaginative ways should be sought to teach it so that doctors become active investigators of their own behaviour. One suggestion to increase active participation was to encourage doctors to develop their own decision aids or clinical scenarios based on their personal and professional experiences. Another respondent suggested that decision making training should be delivered using a range of multiple methods including experiential learning, self-directed learning and some didactic teaching to get the main issues across.

Advice on where to place training: Six respondents gave suggestions about where to integrate clinical reasoning training in medical education. Four people believed it should be included explicitly at the undergraduate level as standard training. Whereas two other people suggested that training in clinical reasoning should coincide with the clinical phases of medical training and would be of most benefit at the postgraduate level, i.e. part of specialised rather than general training. One respondent identified that clinical reasoning training should be linked with courses in evidence based medicine, although did not state why and another was sceptical that it would be successfully integrated in medical education at all due to emphasis on other subjects.

3.5. Discussion

This study reports the results of a questionnaire study completed by medical educators involved in teaching doctors about or an aspect of clinical reasoning. Most courses took place somewhere in the USA and/or Canada. The vast majority of training was delivered in a university. The content of courses usually included a focus on helping doctors make better clinical judgements or decisions and/or developing skills in patient centred care or shared decision making. Most courses were informed by more than one type of theoretical framework, usually classical decision making theories and the information processing approach. There was also reference to more than one type of application of decision theory such as risk communication, decision analysis or Bayes theorem. Courses that were integrated into medical curricula were mostly delivered as a module on clinical reasoning or a module on a broader topic. With regard to the design of future courses, respondents believed a variety of topics should be taught using different forms of teaching methods. There was a consensus that a course on improving clinical reasoning is more appropriately integrated at the postgraduate level. However, the best time during undergraduate training was considered to be alongside communication skills training and/or during times of clinical attachment. There were mixed views about the likelihood that clinical reasoning training would gain wide support in medical education. These results are discussed in more detail below to answer the study's objectives.

3.5.1. Types of clinical reasoning courses in undergraduate and postgraduate medical education

Two types of clinical reasoning courses in undergraduate and postgraduate medical education were found to be common. One type was enhancing doctors' communication and shared decision making skills and the other type was helping doctors make better clinical judgements and decisions. Courses that focused on improving communication and shared decision making, emphasised the role of the patient in the decision process. This view may exist because shared decision making is advocated as an ideal approach to making clinical decisions (Charles, Gafni, and Whelan 1997; GMC 2009) and/or it reflects the specialist interest of the people approached to volunteer in the study. Other courses that were aimed at helping doctors improve their own clinical reasoning typically drew upon normative decision theory. These courses taught doctors how to convey risk accurately, apply decision analysis and/or Bayes theorem. This finding is in line with Elstein's earlier surveys (Elstein 1981; Elstein et al. 1985) and the systematic review (Kurzenhauser & Hoffrage 2002; Grant et al. 2006; Lee et al. 2007). Learning to communicate risk accurately to patients was considered to be an aspect of good clinical reasoning. This was not emphasised as it was in earlier surveys (Elstein 1981; Elstein et al. 1985). It suggests that risk communication has gained importance in medicine over the years, particularly with increasing interest in shared decision making (Godolphin 2003; Sedgwick & Hall 2003; Edwards, Elwyn, Wood, Atwell, Prior, and Houston 2004).

Overall, these results indicate medical educators still believe in the importance of training doctors to calculate optimum decisions by using applications of normative decision theories. Whilst this approach offers logical ways of resolving difficult decision problems, the practical relevance to the clinical context is likely to be limited to situations when time is plentiful and distractions are minimal (Croskerry 2005). Often the doctor is working under conditions that would not permit enough time to allow decisions to be calculated using decision trees or Bayes statistical formula. Little attention has been given to investigating how doctors can be trained to make decisions better with these factors in mind.

The systematic review suggested that teaching doctors to be aware of how they reason about decisions is associated with better clinical decision making, at least hypothetically. For example, participants trained to be aware of heuristic thinking and cognitive errors in the clinical reasoning process developed enhanced decision making skills (Wolf et al. 1988; Hershberger et al. 1995; Round 1999). In this study, heuristics and bias was identified by most people as a topic that should be taught in a course seeking to improved doctors' clinical reasoning. In practice, few respondents said they included this in their training courses. Whilst nearly two thirds of respondents stated they referred to the information processing approach, there was little evidence that improving the way doctors reason about clinical problems was important.

3.5.2. Inclusion of courses into medical curricula

The majority of courses had been integrated into medical curricula. This finding differed from that of the systematic review (Chapter 2) where the majority of courses had been designed as one off studies that were not included in medical curricula. Together, the results from both studies demonstrate that training doctors improve their clinical reasoning has not gained importance in medical education. It was unclear as to whether courses were temporarily included in curricula as a trial or became a permanent feature. Some of the courses were delivered as modules on clinical reasoning, but mostly aspects of clinical reasoning were included in a communication skills module. This suggests there is a view that communication and making judgements and decisions are interrelated in clinical practice (Windish et al. 2005).

There was a shared belief that the success of clinical reasoning training depended on a sufficient amount of clinical experience. Most respondents believed that clinical reasoning training should be introduced at the postgraduate rather than undergraduate level.

There was concern that undergraduate students would not recognise its relevance due to their limited experience of making judgements and decisions in the clinical context. This is contrary to an alternative belief that the early inclusion of formal training in clinical reasoning during undergraduate years is necessary (Elstein et al. 1978; Elstein 1981; Elstein et al. 1985; Round 1999; Croskerry 2005; Croskerry 2009b) and contradicts evidence suggesting that medical students without any clinical experience can benefit from formal clinical reasoning training (Margolis et al. 1982; Vollebregt et al. 2005). These findings are encouraging for educators who are keen to introduce training in clinical reasoning early on in medical education to students with little experience in practice.

Most people held ambivalent attitudes towards the impact of clinical reasoning training in medical education. There was a disparity between the perceived importance of this type of training and its practical implementation in medical education. Whilst educators in this study believed that improving clinical reasoning would help to increase patient safety, many also believed that training in this area would not gain support in medical education if others in senior positions would not prioritise it. Currently this perception is a barrier to the implementation of clinical reasoning training in medical education (Taut & Alkin 2003). In response to another questionnaire, all respondents believed that making clinical judgements and decisions was very important to their practice, but the overwhelming majority had not received any formal training or read material explicitly on this area (Croskerry 2005).

3.5.3. Views and preferences of how clinical reasoning courses should be designed

There was an agreement that the content of future clinical reasoning courses should include a range of topics and that a range of teaching resources should be used to deliver courses. Participants thought the breadth of topics should include risk perception, decision analysis, diagnostic test accuracy, clinical reasoning, heuristics and bias, practice guidelines and cost-effectiveness analysis. Elstein et al

(1985) also found similar topics were also considered important in future clinical reasoning courses. This further supports the traditional view that doctors should be trained to use techniques that improve the outcomes of their judgements and decisions, for example an accurate diagnosis. The teaching of clinical reasoning methods and heuristics and bias was considered important by most respondents, but very few courses included these topics. This indicates a gap between what is perceived to be important to teach trainees and what is actually being taught.

It is unlikely that it would be feasible to cover a range of diverse topics in one session or a short course about clinical reasoning. Further, the inclusion of too many topics in one course may not necessarily be ideal. In the systematic review, most interventions focused on teaching one aspect of improving clinical reasoning, such as decision analysis (Margolis et al. 1982; Lee et al. 2007) Bayes theory, (Kurzenhauser & Hoffrage 2002; Grant et al. 2006) or heuristics and bias (Wolf et al. 1988; Hershberger et al. 1995; Round 1999). Only one intervention delivered a wide range of topics (diagnostic test accuracy, decision analysis, cost-effective analysis, critical appraisal of literature) and this was not well received by medical students, who found the volume of information to be too much (Cebul et al. 1984). Participants' reactions towards a training programme influences whether they go onto apply their learning (Noe & Schmitt 1986; Kirkpatrick 1998). It is important that any clinical reasoning course is well received in the first instance.

3.5.4. Implications of findings to medical education

If improving doctors' clinical reasoning is to gain importance in medical education, there needs to be a shared understanding amongst medical educators about how to improve doctors' clinical reasoning. Amongst the courses that had been designed, there was little shared understanding about how to improve doctors' clinical reasoning. Further, for this training to be effective the design of future courses should be informed by the evidence reviewed in Chapter 2.

The results of the systematic review and this study indicate that educators prioritise the improvement of decision accuracy (outcome) rather than the quality of reasoning about clinical problems (process). Teaching doctors how to reason well should not be overlooked, particularly as the effectiveness of this approach is supported by evidence (Wolf et al. 1988; Hershberger et al. 1995; Round 1999; Abraham et al. 2004). The value of this approach should be promoted amongst medical educators as another way to improve the quality of doctors' clinical judgements and decisions.

3.5.5. Critique of study

This study has some limitations. Whilst effort was made to seek out the main medical decision making related groups for distribution of the questionnaire, the study sample may not be truly representative. It is possible that some educators chose not to participate or they were unaware of the study, and so their course details and views are not represented here. The responses to questions 7 and 8 did not discriminate between a most important and least important topic and method of teaching to use. The majority of people felt most or all topics that were listed should be included in a clinical reasoning course. This was a limitation of the wording as respondents were instructed to tick all options that apply. Perhaps it would have been more informative to allow respondents to select one option only or to rate each option in order of importance.

The strengths of this study are that it builds upon previous work. It updates the findings of earlier surveys (Elstein 1981; Elstein et al. 1985) and provides more detailed information about the types of clinical reasoning courses that have been developed. The data represents information from a wider range of medical educators, not just those associated with the Society for Medical Decision Making. This study supplements the findings of the systematic review (Chapter 2) by integrating unpublished information about training doctors in clinical reasoning. The results of this study with the systematic review provide an evidence base that should be used to guide the design of future courses to help doctors improve their clinical judgements and decisions.

3.6. Summary

The results of the systematic review and this study show that not many courses to improve doctors' clinical reasoning have been developed. Even fewer have managed to gain inclusion into medical curricula. This indicates that clinical reasoning training has not gained importance in medical education. There is little agreement between educators about how to enhance doctors' clinical reasoning effectively. However, training to improve the outcome of a judgement and/or decision remains a more common approach than training to improve the process.

4.

Doctors' perceptions of making clinical judgements and decisions effectively: an interview study

This chapter describes a qualitative survey carried out to explore the perceptions and experiences doctors have when making clinical judgements and decisions. It first reviews the qualitative literature that has explored doctors' clinical reasoning and states the aim and objectives of the current study. The methods are described, including the rationale for using interviews for data collection, development of an interview schedule as well as a coding frame to analyse the data using a thematic framework approach. The themes generated from the analysis are discussed and presented according to the study's objectives. The chapter ends with a discussion about the findings in relation to existing literature and the implications to medical education and practice.

4.1. Background

In chapters 2 and 3, the existing evidence base of evaluated and non-evaluated clinical reasoning interventions was identified. The present study was concerned with gaining in depth information about doctors' perceptions and experiences about how they make clinical judgements and decisions. Research attempting to understand the nature of clinical reasoning methods began over thirty years ago. Initially, evidence from experimental studies suggested that diagnosis was a process of hypothesis generation and testing (Elstein et al.

1978). This was soon challenged in favour of a more automatic process that involved categorising and recognising clinical signs and symptoms (Patel & Groen 1986; Barrows & Feltovich 1987). More recently, the complexity of doctors reasoning about clinical problems has been recognised. Authors have now suggested that doctors draw upon a repertoire of different reasoning methods depending on the nature of the clinical problem (Elstein & Schwartz 2002; Croskerry 2002; Norman 2005). These different reasoning methods have been categorised broadly into two main types; a heuristic and systematic processing method (Chaiken 1980; Chen & Chaiken 1999; Payne & Bettman 2004; Croskerry 2009b) (Section 1.5.1).

A review of doctors' clinical reasoning in the qualitative literature was conducted. Most studies focused on gaining deeper insight into how doctors experience and manage uncertainty during decision making, or on identifying factors that influence their clinical judgements and decisions (Bendtsen, Hensing, Ebeling, and Schedin 1999; Lockey & Hardern 2001; Grant & Dowell 2002; Fauriel, Moutel, Duchange, Montuclard, Moutard, Pierre Cochat, and Herve 2005; Farnan, Johnson, Meltzer, Humphrey, and Arora 2008). A few studies were identified that had explored doctors reasoning processes using qualitative methods. In brief, Coderre et al. (2003) explored the relationship between diagnostic reasoning strategies and diagnostic success. A think aloud method revealed doctors employed three types of reasoning strategies when making diagnoses; pattern recognition, schema-inductive and hypothetico-deductive method. Diagnostic judgements were more likely to be correct when doctors had used pattern recognition or schema-inductive method. The hypothetico-deductive method was the least associated with diagnostic success. Two studies have described the types of reasoning processes doctors used when deciding how to manage specific conditions (Denig, Witteman, and Schouten 2002; Jacklin, Sevdalis, Darzi, and Vincent 2008). General Practitioners were required to think aloud when making decisions for scenarios of patients with urinary tract infections and stomach complaints. Results indicated that the clinical decision process resembled a heuristic rather than systematic approach. Doctors did not always reach decisions by evaluating all known options, rather, many were based on routine prescribing habits (Denig et al. 2002). Surgeons were asked to describe the most important decisions they had to make when operating on patients with gallstones (Jacklin et al. 2008). On average 18 decisions were important. Two decision making strategies were described, one being an intuitive method based on doctors' clinical experience and another was the use of specific personal decision rules. Both resembled a heuristic approach to decision making. The researchers noted that the type of strategy employed depended on the decision problem. Doctors were more inclined to rely on their intuition and past experiences when considering whether to operate or not. Whereas decisions in theatre tended to be more systematic. Doctors used personal decision rules that resembled an 'if-then' approach. For instance, *if* the gall bladder is perforated *then* it should be extracted (Jacklin et al. 2008).

The findings above demonstrated doctors' use of heuristic and/or systematic cognitive processes to reach clinical judgements and decisions. These qualitative studies corroborate the findings from previous experimental studies. The same types of clinical reasoning methods were identified and the idea that doctors draw on a repertoire of various reasoning processes was reinforced. Overall the quantitative and qualitative literature has furthered our understanding of the cognitive processes doctors use to make clinical judgements and decisions. However, a limitation with previous experimental studies is an emphasis on identifying diagnostic reasoning strategies. This has led to better understanding about how doctors' think about diagnosis compared to their reasoning about the management of patients (Norman 2005). A limitation of the qualitative studies is that clinical reasoning processes were studied in relation to specific conditions in specific clinical contexts. While this provides in depth information about how doctors solve particular clinical problems, it does not capture the complexity of clinical reasoning. The present study built on this qualitative evidence from a broader approach. Doctors from a range of clinical specialties were included so that discussions were not restricted to decision making about specific conditions in one clinical context. The study also focused on understanding the processes used to make diagnoses and patient treatment decisions.

4.2. Aims and objectives

The aim of the interview study was to explore doctors' views and experiences of how to make clinical judgements and decisions effectively. The research objectives were to:

- identify factors that influenced doctors' clinical reasoning about clinical cases;
- describe doctors' perceptions of an ideal or 'gold standard' way of making effective clinical judgements and decisions;
- describe the strategies doctors report they use to make clinical judgements and decisions;
- explore doctors' views about how to develop clinical reasoning competency.

4.3. Method

4.3.1. Design

This study used a cross-sectional survey design employing qualitative methods. A semi-structured interview technique was used to collect data. There are circumstances when a qualitative method of study is more suitable than quantitative methods (Coolican 1999; Willig 2001; May 2002; Robson 2002; Smith 2008). These include studies that focus on describing perceptions and experiences of a particular phenomenon (Coolican 1999; Willig 2001; May 2002; Robson 2002; Smith 2008), an in depth exploration of an under researched area, and making use of a rich data source to inform the design of future quantitative studies (King 1994). The present study focused on exploring a range of doctors' views and experiences of making clinical judgements and decisions, therefore the choice of a qualitative method of study was appropriate. It also sought to provide a rich and detailed description of the learning of clinical reasoning skills and methods doctors believe they use to make judgements and decisions. In addition, it was anticipated that the rich data source may be used to inform the design of an intervention to improve clinical reasoning and to make recommendations for future training purposes.

For this study, interviews were the preferred method of data collection. Interviews are advantageous because they enable interaction with participants in a way that questionnaires or observations did not permit. This creates opportunity for the researcher to develop rapport with participants and gain detailed information about the topic of interest. Interviews offer a flexible method of data collection, the researcher is able to clarify participants' comments and follow up interesting ideas. This is especially useful for exploring a poorly understood topic in detail (Robson 2002). However, an interview method is not without limitations. The presence of an interviewer may introduce bias into the data collected, in the following ways. First, participants may feel inclined to respond to questions in a way they think is expected of them – a phenomenon known as social desirability bias. Second, the researcher's prior knowledge, beliefs and assumptions about the discussion topic can affect the way the data is interpreted at the analysis stage (May 2002; Robson 2002). The interviewer can reduce the chances of socially desirable responses by wording questions neutrally and refraining from sharing their opinions during the discussion. To reduce the researcher's subjective interpretations of the data, interviews can be audio recorded so there is an accurate, permanent record of the entire discussion. The data can also be coded using participants' own words and phrases at the analysis stage. As with most research designs, it is possible that there are differences in the type of people that volunteer for interviews compared to those who do not. This means the data generated may not represent the views and experiences of the entire population of interest. However, qualitative research is more concerned with gaining in depth information about the views and experiences of the study sample and less concerned with making broader generalisations (Willig 2001).

Interviews can be conducted over the telephone, in a group setting and/or individual face to face (Coolican 1999; Robson 2002; Fontana & Frey 2005). Telephone interviews may have been more convenient than face to face interviews or encouraged greater participation. However, the less interactive nature of the discussion meant that rapport was less likely to develop between the author and participant. Further, the data may have been difficult to interpret accurately if participants had not been physically present during the discussion (Robson 2002). An alternative may have been to hold a focus group discussion. This is an efficient way of gaining a large amount and range of data from several people at the same time. The group interaction enables shared and opposing views to be identified with ease (Robson 2002). However, group interviews are difficult for researchers to control (Robson 2002; Fontana & Frey 2005). The traditional question-then-answer format may have been more difficult to maintain in a group setting as participants speak over one another. Also the personalities and views of some members might have dominated over others in the group (Robson 2002; Fontana & Frey 2005). There is also a lack of anonymity and confidentiality in a group setting which may discourage some people from participating. Interviews that are conducted individually and face to face with the researcher do not have the above disadvantages. It was particularly important that good rapport was established so that participants spoke frankly about their experiences of making judgements and decisions in practice. Individual face to face interviews were more likely to ensure a feeling of anonymity and confidentiality than telephone or group interviews (May 2002). Consequently this was the chosen method of interview format.

Interviews can be conducted from several different approaches that range from structured, unstructured or semi-structured. In a structured interview, each participant is asked the same questions in the same order with a choice of pre-selected answers (Robson 2002;

Fontana & Frey 2005). This approach is similar to a participant self completing a questionnaire, except it is completed by an interviewer. This can be useful if a study aims to reliably compare responses of different participants and possess a method that is easily replicated by others (Coolican 1999). However, this method would not have allowed the author to engage in natural conversation with participants, and consequently the data collected would have lacked detail (Coolican 1999; Fontana & Frey 2005). In contrast, unstructured interviews have little or no standardisation between interviews. Researchers do not necessarily need to have planned specific questions to ask, rather general topic areas guide the discussion (Coolican 1999; Robson 2002). Consequently participants have a lot of control over the topics of conversation and are free to speak about what they like. This approach has the advantage of encouraging an interactive discussion between researcher and participant that resembles a natural conversation. However, as the present study had a defined agenda to meet, a more structured interview approach was necessary (Robson 2002).

An alternative approach that offered a combination of structure with flexibility was the semi-structured interview (Ritchie & Lewis 2003). A sufficient balance of control over the interview is achieved for both researcher and participant. Like a structured interview, the researcher can use the interview schedule as a guide rather than a script to ensure the data generated answered the research questions appropriately (Robson 2002; Fontana & Frey 2005). Questions are still pre-determined to ensure the study objectives are fulfilled, but the wording and/or order of questions may be changed to suit individual interviews. The researcher can also use a range of probing techniques to ensure important issues raised spontaneously by the participant are followed up in enough depth (Ritchie & Lewis 2003). At the same time the participant is allowed enough scope to share the views and experiences of their choice and the discussion still resembles a relaxed and informal conversation (Willig 2001; Robson 2002). Overall, it was considered that a semi-structured approach to interviewing combined the advantages of structured and unstructured interviewing approaches. Participants of the present study were interviewed face to face using a semi-structured approach because of the advantages of structure and flexibility.

4.3.2. Ethical approval

The study received ethical approval in April 2009 from the Leeds (East) NHS Research Ethics Committee (see Appendix 8.6).

4.3.3. Sample size

Within flexible research designs it is difficult to pre-specify how many participants are required. It is acceptable that the researcher makes a judgement as to when data collection has reached saturation (Robson, 2002). This point is reached when the researcher feels that further data will not contribute anything new to the current data set. Factors such as the breadth and nature of the research topic and the data collection method used are likely to influence the number of participants required to reach a level of saturation with the data (Morse 1995). Some authors have estimated that a minimum of 6 interviews are necessary to reach saturation and a maximum of somewhere between 20 and 30 interviews should be conducted (Morse 1995; Creswell 1998; Guest, Bunce, and Johnson 2006). Guest et al. (2006) reported that although 60 interviews had been conducted, saturation had occurred within the first 12 interviews. Based on this guidance it was estimated that between 6 and 20 interviews were sufficient for the present study.

4.3.4. Sampling method

Participants were selected using a purposive sampling recruitment method. Purposive sampling is suitable in flexible research methods, particularly when the sample needs to reflect the views from a range of individuals (Robson 2002). In the present study, it was important to represent a range of doctors' views according to different clinical specialty and expertise level. Therefore a purposive sampling method permitted the researcher to deliberately recruit participants with these objectives in mind. A purposive sample is one type of non-probability sampling method that does not make any statistical inferences about the likelihood of selecting individuals (Robson 2002). For this reason, a limitation of selecting participants purposefully is that the findings may not represent the views and experiences of all doctors. However the importance of fulfilling the study objectives outweighed this limitation.

The author used contacts from the School of Medicine to approach doctors who worked at the Leeds teaching hospitals and Leeds Institute of Health Sciences. These people were initially invited to participate in the study via e-mail. On behalf of the author, the School of Medicine's Director of Student Support sent a mass e-mail to fourth and fifth year medical undergraduates. Those that responded with interest were sent a study information sheet electronically. Paper copies of the information sheet were also distributed to postgraduate students studying for the Certificate in Health Research. This particular course was targeted as a proportion of the students were known to be medical doctors. Further participants were sought from the contacts of individuals that had agreed to take part in the study.

4.3.5. Sample characteristics

To represent a range of views and experiences about decision making, the sample included doctors that worked in a variety of clinical specialities and who were at different stages of seniority. The sample included doctors that worked in primary care and a range of medical and surgical secondary care specialties. Specifically, the sample consisted of (i) general practitioners so that judgement and decision making around a wide range of acute, chronic, mild and severe illness was represented, as well as decisions to refer patients to specialists, (ii) an anaesthetist, obstetrician and paediatrician for their experience of making decisions under particularly high risk and emergency conditions, (iii) a nephrologist and psychiatrist for their experience of making judgements and decisions about long term conditions with kidney and mental health patients and (iv) a vascular and paediatric surgeon for their experience of making decisions in theatre.

The sample also included a range of seniority levels from novice medical students, newly qualified doctors and consultants at the most experienced level. There is evidence to suggest that doctors of different levels of clinical experience make decisions differently (Jensen, Shepard, and Hack 1990; Croskerry 2002; Schmidt & Rikers 2007). The sample included undergraduate medical students in their fourth and fifth year of study, newly qualified postgraduate FY1 and FY2 doctors, specialist trainees and consultants. Undergraduate medical students below fourth year were not invited as they would have little experience of making judgements and decisions in practice due to less exposure to clinical settings. For analysis purposes, the sample was divided into three broad categories to reflect doctors' different levels of seniority. Junior doctors were classed as medical students and postgraduates in their foundation years of training (FY1 and FY2's). Middle grade doctors were general practitioners with less than five years of general practice experience and specialist trainees. Senior doctors were general practitioners with more than five years of general practice experience and those who had reached consultant level.

4.3.6. Materials

The study's materials included an information sheet, consent form, interview schedule and a coding frame for the analysis. The details of each material are described further below.

4.3.6.1. Study information sheet

A study information sheet was designed according to the NHS ethics committee and university ethics committee guidelines (Appendix 8.7). The information sheet explained the purpose of the study, why the individual had been invited to take part, what the study involved and how the data would be used and stored. Contact details of the author were provided and it was explicitly stated that participation was voluntary, their identities would be hidden and the data they generated would be used strictly for research purposes only.

4.3.6.2. Consent form

A consent form was developed according to the standard format of the Leeds NHS Research Ethics Committee (Appendix 8.8). Informed consent was obtained upon completion of this form. Participants were required to indicate they had read each statement and were required to sign and date the form.

4.3.6.3. Interview schedule

The content and format of the interview schedule was developed using good practice guidelines in qualitative methods (Coolican 1999; Robson 2002; Ritchie & Lewis 2003). Questions were generated based on a review of the clinical reasoning literature and in accordance with fulfilling the study's objectives. To achieve good rapport with participants and data of sufficient breadth and depth, a mixture of open and closed questions were included. Open ended questions were formed from the key topic areas in order to explore answers in a flexible and deep manner (Ritchie & Lewis 2003). Key areas were related to challenges at the workplace, good and poor clinical reasoning, factors that may affect judgement and decision making and learning to acquire clinical reasoning skills. Prompts were also prepared in anticipation to ask participants to expand on interesting comments and/or provide them with clarification for some of the broad questions. Prompts and follow ups were phrased as open ended, but on occasions as closed, questions and this proved a useful way of maintaining focused and detailed discussions (Ritchie & Lewis 2003). Effort was made to word questions in short sentences using clear and neutral language to encourage participants to speak frankly (Coolican 1999; Robson 2002; Ritchie & Lewis 2003). The interview schedule was structured in a sequence considered to be good practice in qualitative research. Straight forward questions were asked at the beginning to ease participants into the interview and create some rapport with the author (Robson 2002; Ritchie & Lewis 2003). Participants were asked to introduce themselves in terms of their seniority level and the clinical specialty they currently worked in. The interview then moved on to the main body of questions that represented the purpose of the interviews. Questions in the main body were divided into two main areas related to the experience of making clinical judgements and decisions in practice and the development of clinical reasoning skills. When answering these main questions, participants were asked to describe examples from their clinical experience and formal medical education to illustrate clearly their perceptions and experiences. Interviews drew to a close with the author signalling which question would be the last and ended with inviting participants to say anything they wished to add to the discussion (Robson 2002; Ritchie & Lewis 2003) (see Appendix 8.9 for interview schedule).

4.3.6.4. Coding frame

A coding frame was developed in order to classify the transcript data by categorising the texts into categories and themes (Appendix 8.10). The development of the coding frame is described in more detail in the analysis section 4.3.7. Thematic framework analysis method was used to inform the design of the coding frame (Ritchie & Lewis 2003). The coding frame organised the transcript data into the following six themes: (i) perceptions of clinical specialty, (ii) factors which influenced reasoning, (iii) perceptions of making judgements and decisions, (iv) novice and expert,(v) professional practice, (vi) role of learning. Each theme was associated with five or six categories.

4.3.6.5. Pilot study

A pilot interview was conducted to evaluate the suitability of the interview schedule and the author's interviewing skills. It is important to check that a researcher's style of interviewing and questions asked obtain a full and coherent account of the main topics (Robson 2002; Ritchie & Lewis 2003). Particular attention was given to assessing the suitability of the sequence of questions, the use of language and ease of understanding questions. This also gave some prediction of the duration other interviews may take.

The pilot interview was conducted with a consultant obstetrician who was selected due to his senior level of clinical experience and research experience at the university. The pilot interview followed the format of the interview schedule, the participant answered a question and then gave feedback about the nature of that question until the interview reached the end. For each question the participant was asked to comment on its clarity, appropriateness and whether they had any suggestions for improving the interview schedule. This feedback was also used to determine whether the study's objectives had been fulfilled. The audio recording was listened to and transcribed verbatim by the author. The transcript and the author's impressions of the pilot interview were then discussed with supervisors.

As a result of the pilot interview, some problems were identified. The interview schedule had not generated data of sufficient clarity, scope or depth. It was decided that the interview schedule should be revised in its entirety. The main problems were that questions were unclear and too broad and this caused the author to over prompt the participant. Consequently the data elicited was unfocused and did not sufficiently fulfil the study's objectives. On reflection questions about abstract concepts were worded in an overly direct manner, for example 'Can you talk me through the process of how you make a clinical decision?'. This proved too challenging for the participant to describe clearly and it highlighted the difficulty others may face when trying to articulate their thought processes. To combat this, questions about how decisions are made and how this skill is developed, questions were phrased in a more implicit but specific manner. This was achieved by asking participants to narrate examples when answering each question to illustrate their points more clearly. Examples were 'Can you think of an example of a patient who had a good medical outcome but you felt uncomfortable with an aspect of your clinical judgment?' and 'Are there any examples of when you would have made a decision differently?'. Such questions were answered with ease and enough detail and the author was able to gain an implicit insight into participants' decision making processes. The emphasis on participants' providing examples from their clinical and formal education experiences became the key technique to achieve coherent and meaningful discussions. This worked well and was implemented in every subsequent interview.

4.3.6.6. Procedure

Interviews were organised at a mutually convenient time with those who agreed to take part, usually in a quiet setting at a participant's place of work or study. This included various departments around Leeds teaching hospitals and the University of Leeds. Participants were recruited until no new themes of data emerged and a sufficient number of doctors of varying seniority levels and clinical specialties had been interviewed.

A few days prior to each interview, those that agreed to participate were asked to think about the following questions from the interview guide in preparation;

> Can you think of an example of a patient who had a bad outcome but you felt comfortable with your clinical judgment?
>
> Can you think of an example of a patient who had a good outcome but you felt uncomfortable with an aspect of your clinical judgment?
>
> Are there any examples of when you would have made a decision differently?

These questions might have been difficult for participants to answer spontaneously during the interview. As they were expected to generate an insight into the processes doctors use when making judgement and decisions, these questions represented a key part of the interview. Therefore it was considered appropriate to ask participants to prepare their answers to these questions in order to avoid situations of not being able to think of examples during the interview. This procedure worked well as all participants came to their interview with examples they had prepared in advance, which led to stimulating discussions.

At the beginning of each interview, participants were asked to spend a few minutes reading a paper version of the information sheet and then complete and sign the consent form. The interview always began with a discussion of participants' stage of medical career and their current clinical specialty of work. This background information was helpful in guiding the rest of the discussion. For most interviews, questions were asked in the order they appeared on the interview schedule and prompts were added spontaneously as the discussion progressed. Also, on a few occasions, the order of questions was changed spontaneously in response to the direction of the discussion, but all questions were asked at some point. This helped maintain a conversational manner. On completion of the interview schedule, participants were asked whether they would like to add anything else to the discussion. Interviews ended with a debriefing of how participants felt about the interview, if they had any questions and whether they were happy for the audio-recording to be analysed. Each person was thanked for their time and invited to contact the author in the event of any issues arising. On average, interviews lasted 45 minutes but ranged from 25 to 60 minutes.

Immediately after each interview, the author listened to the recording and noted down her initial thoughts and impressions of how the discussion went and the main themes that emerged. These notes became memos that formed a preliminary analysis of the data. Each recording was copied for back-up purposes and transcribed by a third party. The qualitative data analysis package NVivo (QSR international, version 8) was used to organise and store the data during the analysis process.

There were several methods adopted to increase the reliability of the study. This included the use of a good quality audio-recording device, a professional transcriber and the qualitative data analysis software, NVivo. The interviews were audio-recorded for several reasons. Firstly, it eliminated the need to take notes during interviews so the author was able to devote full attention to the discussions (Coolican 1999; Ritchie & Lewis 2003). Secondly it provided a verbatim permanent record of each discussion from beginning to end. This helped increase reliability as the author could use feedback received between each interview to improve her interviewing skills (Robson 2002). Also, verbatim transcriptions of the recordings enhanced the reliability of the data analysis process as each transcript was analysed in full. This meant that all of the data rather than a selection of data were closely examined. This reduced the chances of a biased interpretation of the data based on the author's prior assumptions (Ritchie & Lewis 2003). A potential disadvantage of audio-recorded interviews is that some individuals may feel inhibited due to anonymity and confidentiality reasons (Ritchie & Lewis 2003). This was not perceived to be problematic during the present study as the topic of discussions were not of a sensitive nature. The author explained the value of audio-recording and reassured participants that she would be vigilant to conceal any identifying information. Procedures to maintain confidentiality were also clearly explained in terms of how the recordings and transcripts would be stored and used. Each interview was transcribed by an experienced, professional transcriber using a word processor. Interview recordings were transcribed verbatim to produce an authentic electronic version of the discussions, including expressions pauses and laughter. Transcriptions were checked by the author for accuracy to further increase reliability. Finally, NVivo software provided an efficient filing system to navigate through and organise a large textual data set.

4.3.7. Analysis

The aim of the analysis was to accurately describe participants' views and experiences of making effective clinical judgements and clinical decisions. This was achieved by summarising the data set in the form of themes with related subcategories, using a thematic framework method (Ritchie & Lewis 2003; Braun & Clarke 2006). A thematic framework method was suitable as it generates themes and categories from the data itself. This was important as the focus of the present study was to illustrate the participants' own perceptions and experiences of making clinical judgements and decisions. Other methods of qualitative data analysis were considered such as Grounded Theory and Interpretative Phenomenological Analysis. Grounded Theory aims to discover a theory about a particular experience that emerges from the data (Glaser & Strauss 1967). However, as Grounded Theory is not a descriptive method of qualitative data analysis, this was not selected as the aim of the present study was to describe the data, rather than generate a theory that explains people's views and behaviours (Coolican 1999). The central focus of Interpretative Phenomenological Analysis is to provide detailed insight of how people experience and make sense of particular phenomena (Smith & Eatough 2007). Its idiographic emphasis encourages the study of small homogenous groups and it is necessary to achieve a balance of phenomenological description with insightful interpretation (Smith & Eatough 2007). To meet the objectives of the present study, it was necessary that the sample consisted of range of doctors to represent a range of views. Further, this study intended to be descriptive rather than interpretive of the meaning behind doctors' views and experiences of clinical judgement and decision making.

The stages of thematic framework analysis are easy to follow and encourage the author to systematically and rigorously attend to the entire data set. However there is still room for a degree of flexibility when following these key stages (Braun & Clarke 2006). At the centre of the thematic framework method is the development of a coding frame i.e. a thematic framework. Its purpose is to classify and organise the data into themes and related categories (Ritchie & Lewis 2003; Braun & Clarke 2006). The following six steps were taken to develop a reliable coding frame. Steps one to four were carried out manually using paper versions of transcripts, whilst steps five and six were managed by Nvivo software.

4.3.7.1. Identification of meaningful units

In order to gain an overview of the breadth and depth of the data all transcripts were read through once. This was a useful way to become familiar with the complete data set and the possible patterns. Five transcripts were selected to be analysed in more detail. Meaningful phrases or sentences were identified and used to divide the texts into units. These units represented the different thoughts and ideas of participants. The example below illustrates how a paragraph from the original text was divided into separate meaningful units:

> Well that's the whole beauty and the problem with primary care and general practice, you have no idea what on earth is going to present.
>
> You could have a number of relatively minor illnesses.
>
> I mean you've got to remember that we're a very busy, really quite deprived area and so therefore say I'd be a duty doc today, I might have 50 people I'll see in a day.
>
> And even on a quiet day if I'm doing a full day it would be a minimum of [sigh] 40 people, 35 to 40 people in one day.
>
> So I mean it can be anything from sort of simple cough/colds, to people coming with anxiety, depression, psychological illness, muscular skeletal problems. Or it could be sort of anything really, cardiovascular, respiratory...
>
> ...so it's a massive range, a massive range. (Int. 2)

During this early stage of analysis, a selection of transcripts were used to identify meaningful units of data (Ritchie & Lewis 2003). Five transcripts were chosen (interviews 2, 5, 7, 10, 12) that represented views from doctors who differed according to clinical specialty and seniority level, so a range of texts were used.

4.3.7.2. Generation of initial codes

The next stage was to attribute meaning to each unit of data. The author reflected upon the units and decided what codes to assign to each unit. These codes represented a label that summarised the content of each unit according to the author's interpretation. For example, the following unit *well that's the whole beauty and the problem with primary care and general practice, you have no idea what on earth is going to present* was coded under 'context of specialty'. One code was assigned next to each unit of data until all five transcripts had been labelled. The author discussed the codes with supervisors in order to determine whether they represented accurate interpretations of the units of data. As a result of this discussion, it was necessary to go back and divide the texts further into units of data and re-interpret some of the codes. This was good practice to increase the reliability of coding the data. These revised codes were then used to index the rest of the data set.

4.3.7.3. Sorting of codes into initial themes

A separate list of all the codes was made to simplify the sorting of codes into initial themes. The list of codes was carefully examined, duplicates were eliminated and related codes were grouped together. This was an iterative process where codes were re-grouped until a coherent set of groups of codes was established. Each group was then labelled with a title that represented the common link between the codes in each group. These were the initial themes indentified and the codes became the sub-categories under each theme. Twelve themes were identified with a number of associated categories that ranged from 9 to 47. Each theme and its related categories were numbered in the following way and this became the initial thematic framework:

1. Demographics
2. Role – academic
3. Role – educator
4. Role – clinician
5. Stage of training
6. Typical day
7. Typical tasks
8. Advantages of job
9. Continuing professional development
10. Other

To examine its suitability as a coding frame, the initial thematic framework was applied to a transcript that had not been used to develop the coding frame.

4.3.7.4. Revision of initial coding frame

It became apparent that the coding frame in its initial form was not suitable for categorising the transcripts. The process of applying it to a transcript was a difficult and cumbersome task. The initial framework and example of a coded transcript were discussed with supervisors and it was agreed that there were too many themes and categories. This caused there to be a lot of overlap between categories which led to difficulty in deciding where to code some of the data. To rectify this, themes and categories were broadened so that more data could be coded using fewer codes. Four of the themes and many of the categories were eliminated by grouping them together under less specific labels. The first revision produced a much broader and simplified coding frame. It consisted of eight themes with a maximum of seven related categories under each theme.

However, as more transcripts were analysed further revisions were made to the framework. A theme that described the demographics of participants, including their various roles and information about specialties, was removed as it did not contribute to important discussion. Instead, demographic information was presented in a summary table (see Appendix 8.11). It also made more sense to collapse two themes into one so that factors influencing decisions and factors influencing judgements were discussed under an over arching theme about factors that influence the clinical reasoning process in general. The names given to some themes were also revised in order to reflect the content more accurately. For example, making decisions became perceptions of making clinical decisions and learning to make decisions became role of learning. The final thematic framework consisted of six themes. Each theme was associated with between 5 and 6 categories, (see Appendix 8.10)

4.3.7.5. Application of revised coding frame

The revised coding frame was then applied to all the transcripts in order to code the full data set. This stage required the author to make a judgement about which parts of the framework applied to each unit of data. Data was coded by labelling each unit of data with the relevant category number (1.2, 1.3, 1.4...) in the margin of the transcript. A small proportion of the data was miscellaneous and coded under 'other' categories. Overall, coding the data with the revised thematic framework, as opposed to its initial format, was much simpler and quicker. The themes and categories served as an appropriate coding system for all the transcripts and this was a sign of its validity (Ritchie & Lewis 2003).

4.3.7.6. Managing data in Nvivo

The final stage involved restructuring the data set so that each coded chunk of text was grouped under its associated category and theme. This was facilitated by using the qualitative data package, Nvivo. Using the tree node and free node functions, the thematic framework was recreated. Tree nodes represented themes and free nodes represented categories. Each coded piece of text was moved to the relevant free node. Eventually each free node (category) consisted of the full range of coded texts that would be used as examples to describe each category.

4.4. Results

A total of 15 doctors took part in this interview study; 8 males and 7 females. Three doctors worked in primary care, 9 worked in medical and/or surgical specialties and 3 were medical students. Six doctors were involved in academic research and 9 had teaching responsibilities. There were 7 senior doctors: consultants or GPs with more than five years of experience, 3 were middle grade doctors; specialist trainees, and 5 were junior doctors; either foundation year postgraduates or medical students.

Overall, six themes were identified from the data: factors that influenced clinical judgements, factors that influenced clinical decisions, judgements and decision making methods, novice and expert clinical reasoning, professionalism and clinical reasoning, and acquiring clinical reasoning skills. Each theme included five or six related categories. In order to display the results succinctly, themes were organised under four headings to reflect the study's objectives. These were (i) factors that influenced clinical reasoning; (ii) perceptions of ideal judgement and decision making; (iii) methods used to make judgements and decisions and (iv) development of clinical reasoning over time.

4.4.1. Factors that influenced doctors' clinical reasoning

In this section, data from themes 1 and 2 are discussed. The data suggest that different factors influenced how doctors approach diagnostic judgements compared to treatment and management decisions. For this reason, they are summarised in separate sections below. Diagnosis was dependent on having the appropriate clinical information. The patient's appearance, time point at which they were seen by a doctor and physical examination, influenced whether an accurate and timely diagnosis was reached. Decisions about how to manage the patient were influenced by a wide range of factors, some of which included patient risk, preferences and pressure at the workplace.

4.4.1.1. Factors that influenced diagnostic judgements

Doctors said their diagnosis starts as soon as they see a patient and assess their physical appearance. One of the first judgements doctors made was *do they look quite well, do they look unwell just generally?* (Int.8). Doctors stated they were looking for signs of physical problems and poor hygiene, as illustrated below:

> So I just went to see her and saw her and I just thought, she doesn't look really very well, she was breathing quite quickly, had a look at her chest x-ray and she had quite a lot of fluid on her lungs. (Int. 15)

and:

> as the patient comes in you're looking how are they're walking, are they still, are they in pain, are they limping, you know, what's their facial features like, do they look happy, unhappy, are they in pain. If it's somebody you don't know, you know, you will be making other assumptions, you will be looking at how they're dressed, you'll be looking at how kempt they are, are they shaved if they're male, you know, do they smell of alcohol, do they have body odour, do they have a cigarette smell around them. (Int. 2)

However, doctors felt this judgement should be under constant review because a patient's appearance can be deceptive. For example:

> you have to bring the bigger picture in as well because some people don't breathe quickly because they're on opiates and some people don't appear distressed because they're very stoical and things like that. (Int. 15)

In addition, doctors spoke of a gradual appearance of clinical signs and symptoms in conditions that worsen over time. So the time point at which the patient was seen affected their ability to reach a timely diagnosis:

> So because symptoms... and we obviously symptoms at an earlier stage, so the guy who I mentioned had a kidney infection, I saw him at sort of early stage, if I'd seen him a week later you would have said, this guy has a raging kidney infection, how did this stupid GP miss that? Well the stupid doctor maybe missed, if he did, because he saw it at an earlier stage and there was the clinical signs weren't fully developed. (Int.3)

A good physical examination was seen as facilitating the process of reaching an accurate diagnosis. A clear example was:

> And then I examined properly and it all fitted with this textbook pleural effusion, so she had some fluid on her lungs. (Int. 15)

Whereas the absence of a physical examination hindered the diagnostic process, as illustrated in one interview:

> he was very difficult to examine and every time you got near him he got very upset, so it was very hard to work out whether his abdomen was actually tender or not. And this is a problem we have with young children. And I think in the end he turned out to be appendicitis and I'd kept worrying about him and kept going back because I just wasn't convinced that he was medical as he'd been managed that far. (Int.8)

and:

> I basically sort of looked at how he was, did a relatively sketchy neurological examination and sort of said, oh I'm sure things will be fine, you know. And then next day he'd actually got worse, he was admitted to hospital. (Int. 2)

4.4.1.2. Factors that influenced clinical decisions

Decisions about how to manage a patient's condition were influenced by risk factors to the patient, individual preferences and pressure at the workplace. Doctors reported weighing up several different types of risk factors to reach a decision including a patient's age, co-morbid health problems and history:

> And you go along to see them and some cases they may be very sick old people where you would question the surgeon's decision to operate. And you have to take a lot of things into account; the patient's comorbidity, the risk... it's basically a risk/benefit thing, weighing it up. (Int. 10)

and:

> Add on the scenario that she's got a previous caesarean section and that makes the decision even more difficult. So where you're saying, well she should be in labour after she's ruptured her membranes, 24 hours later she isn't, and she's got a BMI of 50-something. So you want her to have a vaginal delivery as much as possible. (Int.7)

Doctors recognised that patients affected their management decisions. For example, patient's preferences and their way of life were considered to meet the needs of each individual patient:

> So she was happy, you know, she'd accepted the responsibility that if it was going to rupture it was going to rupture and that was fine and she didn't want surgery for it. (Int 16)

and:

> What else is going on in the patient's life alters the decision-making process that you have. If that person is looking after somebody who is ill or they have co-dependents, again it alters your management plan and the decision that you make. (Int. 2)

However, doctors talked about the tension between delivering care that is appropriate and in line with patients' wishes. There were occasions when patient preferences had little impact on the decisions doctors made, for example:

> And in that case we were looking to sort of optimise his care and we felt that him being a protracted inpatient was not helping his overall treatment plan, and certainly not beneficial to the behaviours that he was displaying. So it was felt that he needed to have his care transferred from hospital into the community services, although he made very sort of clear his wish potentially to harm himself if we were to take that decision. (Int. 11)

The type of doctor-patient relationship could also influence the management plan for patients. If a doctor was particularly fond of a patient there was more personal investment in the decision which influenced the delivery of care. This was illustrated in the texts below:

> On a personal level there are some patients that you really, really become attached to and you get to know really well, and you almost feel that for them you try even harder . . . and there are other people that make decisions like whether he goes to intensive care and things like that, because he needed to go there in the end, so I found myself negotiating harder rather than just accepting decisions. (Int. 5)

and:

> if you really like someone you're going to push for it, you're going to follow it up, you're going to go there yourself, present the patient's card to the radiologist and say look, I really need this done, can you get this done now. If you don't like someone or you've not spent as much time with them you may not think it's as important and you may not make that effort. (Int. 14)

Doctors acknowledged that work pressures affected management decisions. Peer pressure was experienced as either implicit or explicit. Implicit peer pressure was experienced in terms of the expectations doctors have of each other:

> well what my colleagues think, and sometimes that's implicit and unsaid, so I think, what would other people here do, you know, and what would they think if they saw me admitting lots of people to hospital as well, or, what would they think if you know, they saw this boy with worsening cellulitis two days later and they thought, that fool should have, you know, admitted him. (Int. 3)

At other times peer pressure was a more direct experience such as a colleague insisting an action was carried out:

> Because she [nurse] was hassling me I prescribed it, but then I was weighing it up after I prescribed it, do you understand what I mean? I should have done the weighing up first then prescribed, but my mind was really on making sure that I've done it. (Int. 5)

A recurring view was that the time available affected decisions of how to manage patients. When time was constrained, doctors recognised they would sometimes select the quickest management plan in order to save enough time to deal with other patients:

> It sounds awful but it's the way things are – it would have been a hassle for me bouncing him into hospital because you know, several minutes on the telephone, writing a letter, explaining directions means that you fall behind in your surgery another 15 to 20 minutes, or at least sometimes you can. I did yesterday with a more complex case. Which then has a knock on effect on other patients, and your risk with other patients. (Int. 3)

On other occasions, a lack of time was ignored if the patient was thought to be very ill:

> So for a woman I was dealing with yesterday who I thought might have a . . . potentially have a blood clot in her lungs, it just took lots of time to do a careful examination of her and deal with the hospital and get her admitted. And you know, of course I can see there's a counter at the bottom of my screen which tells me how many patients are queuing for me, so I could see that totting up as time went on. But if something is urgent and important then you know, you just deal with it and take the consequences. (Int. 2)

4.4.2. Perceptions of making ideal judgements and decisions

Data from themes 4 and 5 are discussed in this section. There was no specific process that represented the 'gold standard' way to reach a judgement or decision, or one that should be used on every occasion. In theme 4, there was a sense that the quality of judgements and decisions should be judged on the steps taken to reach them rather than the clinical outcome. Making good judgements and decisions was linked with maintaining professionalism at the workplace. Theme 5 summarised how professionalism was maintained when making clinical judgements and decisions. There was a sense that making clinical judgements and decisions safely and in conjunction with the patient, was ideal. This is explained in more detail below.

4.4.2.1. Safety netting

A generally acknowledged ethos was that clinical judgements should be made in line with patient safety. This idea was a recurring theme across all interviews. Participants referred to their *safety netting* procedures that they used when thinking about a possible diagnosis. These were deliberate acts of precaution the doctor used to protect the patient. One method of safety netting was to give patients explicit instructions about what to do if their condition worsened:

> The other thing I did was... well even when I was uncertain I put in what we call safety net, I said to him explicitly and I documented that, you know, if his pain got worse and you know, if it did turn out to be something like a kidney infection he was to come back and see someone else whilst I was away. (Int. 3)

Or sometimes a patient would be referred for further investigation:

> So sometimes you just do have to refer even if you think the patient is exaggerating or something like that. (Int. 4)

For making decisions about how to manage a patient's condition, another method was to start a treatment regime if the cause of the problem was uncertain. Doctors reasoned that it was better to give a potentially unnecessary treatment than leave a potentially fatal condition untreated.

> So lots of people are started on treatment for things that they haven't got because we can't distinguish but then you treat them for that because that's the most important thing to rule out. The other day someone had symptoms that suggested heart attack. I wasn't convinced and my consultant was though, but either way the right thing to do there is to treat them. Later on the test came back to show

> they didn't have a heart attack but they've already had their treatment for it. (Int. 5)

Adverse medical outcomes were recognised as inevitable in medicine. Doctors maintained confidence in their clinical judgments and decisions by adhering to an accepted procedure at work. For example, if a record of their reasoning was kept, appropriate further tests were ordered or a commonly used emergency procedure was implemented, then decisions were classed as *right* or *good* even if the outcome was bad.

> So in a way you can make a good decision and justify it, and you can justify in a note.... So I can make decisions and involve a lot of uncertainty, but quite comfortably, knowing that I've at least followed some sort of accepted procedure which has some sort of... as well for me which has some sort of epistemological basis. (Int. 3)

and:

> So it was a deterioration that you couldn't have foreseen but I'd already made the process into, I'd already arranged to have the appropriate blood tests, arranged to have the appropriate follow up, arranged to have the appropriate extra information gathering, it just so happened on that occasion, you know, things didn't turn out right. (Int. 2)

and:

> but when you don't know the patient at all, know anything about them and you've done the ABC approach and they still die, then you've done everything you can possibly do. (Int. 5)

4.4.2.2. Shared decision making

Involving patients in decisions was mentioned but the nature of their involvement was important. Doctors would involve their patients in the decision making process as a way of encouraging compliance to their recommendations:

> I always try and involve a patient because if you can get them onto your side and see why you're making the decision then I think that always helps with their compliance and concordance. (Int. 4)

However, there were times when a shared approach to decision making was not ideal:

> I mean the patient has to be involved with that decision-making process. Now I say to be involved with it, but they'd be involved with some part of it. There are some bits where you have to have sort of control. (Int. 2)

4.4.3. Methods used to reach clinical judgements and decisions

Data from theme 3 is discussed in this section. A number of methods emerged that participants believed they used to reach clinical judgements and decisions. It was not clear how doctors selected a particular method to use, but there did seem to be different methods used to reach a diagnosis (judgement) compared to a treatment plan (decision). To assess a patient's health or judge the cause of a condition, doctors relied upon their clinical intuition or an ability to recognise signs and symptoms. Treatment decisions were reached via a deliberate problem solving method or habit.

4.4.3.1. Clinical intuition

Clinical intuition was a difficult concept to articulate. It was described as a *sixth sense, gut feeling* or *knowing something is not right.* Participants trusted these intuitive feelings and allowed them to guide their clinical judgements about a diagnosis. Intuition was used to assess the condition of a patient's health:

> I might look from the end of the bed when the . . . you know, somebody might be busy round the bed and I just look through the window and watch them for a minute, which by no means is any clinical examination, but yeah I might think, well that's fine, that's okay, they look okay. (Int. 8)

Or if there was concern about a patient, it would encourage the doctor to search for further diagnostic information:

> So if I have a hunch that something isn't right then I need to look for something . . . for what isn't right. (Int. 5)

and:

> And he was just vomiting a lot and I think we all know something was wrong but nobody really acted on it and I just remember thinking a lot, this isn't right, something's wrong, and asked lots of different people instead of just going to one Registrar. (Int. 15)

Although, intuition was to be used with caution and was not to be over relied on:

> Although without any substantial evidence I won't make any, you know, real big decision on it, does that make sense? (Int. 5)

4.4.3.2. Recognition of signs and symptoms

There was a shared view that often clinical conditions are easily identifiable from a set of classic signs and symptoms. Many diagnostic judgements were based on an ability to recognise the typical characteristics of conditions:

> So people for example who have a two-week period of low mood, a lack of energy, a lack of enjoyment as well as other tick-box symptoms may fulfil the criteria for say a depressive episode. So we have criteria that we would use in our medical diagnosis model. (Int. 11)

Furthermore, doctors with little clinical experience were expected to be able to diagnose common conditions using this recognition strategy:

> There is a pattern so you know, you'd expect even very junior staff to be able to recognise that set of signs and symptoms and the history and make a reasonable diagnosis and know what to do. So I think that's a form of basic pattern recognition. (Int. 8)

However, most participants stressed that it was important to know that the same condition can present differently in different patients. Doctors needed to be able to recognise atypical as well as typical signs and symptoms of a condition, in order to make an accurate diagnosis. This knowledge was also used to drive the search for diagnostic information:

> It's always a danger that people can jump to conclusions just because it seems to fit a pattern. But then alongside that you have a . . . nothing

> ever fits the pattern exactly but it's how much it's deviating from what you're expecting, and that includes both the presentation and then the subsequent course. So I think it's always important to be looking at what doesn't quite fit and asking questions about that and saying, well what else could explain that. (Int. 8)

4.4.3.3. Problem solving method

At other times doctors would use a more deliberate method than intuition or recognition. This involved careful thought about challenging cases, usually around how to manage a patient rather than diagnose. To reach difficult decisions doctors would evaluate the pros and cons of a situation. An explicit example was:

> What's going on in my head is basically like what the surgeon said, it's sort of...you go, this is an elective procedure, i.e. it's not an emergency. It's also cosmetic in that unlike say I don't know, varicose veins or a hernia repair, you may get pain from your varicose veins or pain from your hernia. The scoliosis isn't causing you any pain or anything like that and you don't die from scoliosis. You can die from a hernia if it gets trapped. So it's a cosmetic elective procedure and you're stood there weighing it all up, you're going, it's a cosmetic elective procedure in a child that's quite little, and you're not happy for whatever reason, you're not happy with the ventilation, it's not quite as it should be. And weighing it up against the fact that you've got all those people waiting, parents that have just left you in tears because they've just left their little girl in there. (Int. 10)

Whilst an implicit example was:

> I think in most sort of decisions which there are . . . there's not one right answer but probably several right answers, it's coming to a consensus and I think negotiating your way through that decision. (Int. 11)

Prioritising tasks and completing them in order of importance, was also part of a problem solving approach. This was a strategy used when time was constrained:

> And if patients come with lots of problem you might decide not to tackle everything at once and put things off. Or you might sort of, instead of sort of doing everything in one go you might think, right okay which is

> the most important, sort that one out and then they can come back and we'll sort the next step out and things like that. (Int. 4)

4.4.3.4. Habit

On other occasions a less deliberative method was use to decide how to treat a patient. Participants implied that they don't need to think hard about clinical decisions if there was a known, set treatment regime or they had developed a preferred method of treatment. These decisions became habituated.

> for that and a lot of other very kind of very common conditions there's a very basic and set treatment regime that everyone . . . so everyone gets a salbutamol inhaler, which is used to help prevent asthma attacks and treat them when they happen. (Int. 14)

> So there's a very, very basic emergency decision that you have to deal with and it's almost like a reflex that doesn't involve your higher centres at all, it's like a spinal reflex, a knee jerk response. You just do the airway, breathing, circulation. (Int. 10)

and:

> I have a favourite way of anaesthetising people for that is actually with something called a spinal, like an epidural. (Int. 10)

4.4.4. Development of clinical reasoning competency over time

Theme 4 and 6 included data about the improvement of doctors' clinical reasoning over time. Senior doctors displayed more flexible thinking and better ability to identify the most important pieces of clinical information, than those with less experience. There was a consensus that clinical reasoning improved through practice at the workplace, but that enough knowledge about disease presentation and physiology should be gained first. Doctors with teaching duties believed they had developed an increased awareness about their own clinical reasoning through teaching. None of the participants had received any formal teaching about how to make clinical judgements and decisions.

4.4.4.1. From novice to expert

There was recognition that novice and expert doctors make clinical decisions about treatment in different ways. Novices explicitly relied on guidelines and protocols at the workplace, as exemplified below:

> if you have got a specific diagnosis there's quite often Trust guidelines and NICE guidelines that you can look at and quite often things like the various colleges like the Royal College of Obs and Gynae[1] has like their own set of guidelines for specific gynaecological and obstetric cases. So there's quite a lot of places that you can look in to get sort of advice for sort of how you would go about treating patients. So it's useful to use that. (Int. 12)

and:

> It probably depends on what the kind of protocol is. Yeah, it depends if you're in that kind of environment where it's very . . . where there are lots of guidelines about your work. (Int. 13)

While consultants displayed more flexible thinking and were able to judge when it was appropriate to deviate from a standard procedure:

> She was at risk of having a sudden emergency and despite her having a general outlook that would mean that we would normally not be very aggressive, I thought we should be aggressive and try and resuscitate her if we could. (Int. 9)

When gathering clinical information to make a judgement about diagnosis, experienced doctors were able to do this in a focused and succinct manner. They were more able to distinguish and obtain essential from non-essential clinical information.

> Even taking the history from the patients, it seems to take them forever, and I suppose we've just got it down to a, you know, what we think we need to know, hopefully what we do need to know. And they don't seem to be able to get to the nitty-gritty of it. It's sort of like they're still scattergun approach of like asking all the questions before they get to the relevant bit. (Int. 10)

[1] That is, the Royal College of Obstetricians and Gynaecologists.

Junior doctors expected to be able to develop this skill over time:

> I know it's getting better but I can't sort of read a set of notes and pick out what's important, and I know that's all about becoming an expert and things like that. (Int. 15)

4.4.4.2. Theoretical knowledge

Most people agreed that a sufficient amount of biomedical knowledge was a necessary foundation for making clinical judgements and decisions:

> I think first principles are always part of your armamentaria, you know, you're going back to first principles and thinking how would this disease produce that. So you have to know how the body works and the processes underlying the workings of the human body. (Int. 8)

Medical students shared the opinion that the signs and symptoms of diseases and other conditions is sufficiently taught in medical schools, but that formal teaching of how to manage these conditions was sparse. This left them feeling more prepared for making clinical judgements (diagnoses) than decisions around managing conditions.

> A lot of the teaching is focused in the first couple of years on recognising symptoms and coming up with possible differential diagnosis. While there's actually not a huge amount of teaching on management and management planning for the patient, you're supposed to kind of pick that up on the wards. (Int. 12)

and:

> so looking more towards management and how you're going to institute your management I think is overlooked in medical school. (Int.15)

Whereas others believed that biomedical knowledge played a limited role in developing their clinical reasoning:

> I don't think any amount of book reading or swatting up or theoretical knowledge is going to compensate for that. You need a theoretical basis obviously for your treatments and your management, but I think ultimately it's about taking decisions and taking responsibility for your actions that's important. (Int. 11)

4.4.4.3. Practical experience

There was a sense that learning to make effective clinical judgements and decisions could only be learned from experience.

> I personally don't think there's been a lot from my training, I think it's just acquisition of experience. I really can't put it down to anything more than that. (Int. 9)

and:

> I think that the best way to learn really is to practice and to be put in situations in which you are impelled to make decisions, under supervision of course initially. (Int. 11)

Colleagues were an important source of guidance on how to make appropriate decisions in practice: Also, it was important to learn from the times when poor judgements and decisions were made so that these would not be repeated:

> And you learn to recognise from observing behaviour as to what is acceptable and recognise that where you would do the same thing it is acceptable, or where you wouldn't do the same thing maybe you ought to be doing it. (Int. 7)

and:

> Well it's just true I think, it's human, it's how we're made to function is that if something goes wrong you really remember it. Well a bit like the case that I discussed earlier. Little things like giving... I'd prescribe a drug and then somebody would say, have you checked their renal function and you go to their renal function and their renal function is completely off and you're like, oh my god – and things like that. (Int. 15)

4.4.4.4. Learning from teaching

Doctors that had teaching duties felt that they had become more aware of how they reasoned about their judgements and decisions. It was implied that this was beneficial for teaching purposes but whether it improved their judgements and decisions in practice was not discussed:

> Yes, definitely, teaching changes you, yeah. It makes you sort of question your own decisions. Yeah, because you're well, I may be asked at any minute why I've said that, so I just think about the answer. (Int. 10)

> but the interesting thing is, is that whereas before I would have to justify to myself what I did, very often when you're teaching you're justifying to students what you do, and that does make you explicitly think about the decision-making process in a way that you don't always do as a practitioner. (Int. 2)

and:

> Sometimes it's hard to follow your own thought processes so far as to how you do things. But I guess the reasons you do something are the same it's just that you have to be able to sort of put that into words to explain it to someone else. Whereas normally you don't tend to think about those words I guess. (Int. 4)

4.5. Discussion

Participants' views and experiences of making clinical judgements and decisions were classified into six broad themes and several sub-categories. The data suggested that doctors' diagnostic judgements are influenced by different factors compared to treatment and management decisions. Further, doctors reported using different methods to make a diagnosis compared to patient management decisions. There was no single optimal method to reach a clinical judgement or decision, but the sample emphasised that the quality of their judgements and decisions should be evaluated based on their reasoning processes rather than medical outcomes. There was a consensus that clinical reasoning is developed at the workplace and improves with clinical experience. No one had received any formal teaching about how to make good clinical judgements and decisions, however those who had teaching responsibilities believed that this role had increased their awareness about their own clinical reasoning.

4.5.1. Factors that influenced doctors' reasoning about clinical cases

Many factors have been shown to influence doctors' judgements and decisions but a distinction has been made in the literature between medical and non-medical factors. Most research has focused on understanding the impact of non-medical factors on doctors thinking as they are seen as sources of bias (McKinlay, Potter, and Feldman 1996). These include factors that are not associated with probabilities of disease such as the patient's personality, the doctors speciality or level of experience (McKinlay et al. 1996). Studies have reported that workplace pressures affect doctors' thinking (Lockey & Hardern 2001; Grant & Dowell 2002; Farnan et al. 2008) and the preferential treatment of some patients over others is well documented (Brotrnan, Stern, and Herzog 1984; Mathers, Jones, and Hannay 1995; Hall, Milburn, Roter, and Daltroy 1998; Hall, Horgan, Stein, and Roter 2002; Bellon & Fernandez-Asensio 2002; Haggerty, Tudiver, Brown, Herbert, Ciampi, and Guibert 2005; Krebs, Garrett, and Konrad 2006). In the present study, there was evidence that doctors' judgements about diagnosis were influenced by similar non-medical factors that were not relevant to discovering the cause of a condition or determining how to treat it. These included a patient's physical appearance, a doctor's relationship with the patient and pressures at the workplace.

4.5.2. Perceptions of an ideal or 'gold standard' way of making effective clinical judgements and decisions

Most participants demonstrated they had some perception of ideal judgements and decisions. They were keen to emphasise that it was more important to evaluate the quality of their judgements and decisions based on steps they had taken to reach them, rather than the consequences of those choices. This finding is in line with a debate in the literature about whether decision clinical quality should be based on outcomes or reasoning processes (Bekker 2010). The uncertain nature of medicine means poor judgements and decisions can result in good medical outcomes by chance and good judgements and decisions can result in adverse medical outcomes. For this reason, an increasing number of decision making researchers stress that the reasoning process is a better marker of decision quality than the medical outcome (Frisch & Clemen 1994; Sox 1999; Yates et al. 2003; Baron 2008; Schwartz & Bergus 2008; Bekker 2010). For participants in this study, a good process included three aspects; (i) patient safety must be at the forefront of doctors' thinking, (ii) their actions and reasoning should be explicitly recorded and justified before others if necessary and (iii) decisions about treatment should be made with patients when appropriate. Participants struggled to articulate anything more specific than the above because it was more important that their practice was satisfactory rather than perfect. A good clinical reasoning process has been defined by others in a similarly broad manner. Schwartz & Berguss' (2008) definition included four aspects, one of which was also that it should be justifiable before others. Sox (1999) stated that the process should end with shared decision making. In this study, participants agreed that it was ideal for patients to be involved with clinical decisions. However, there were a range of views about what patient involvement meant.

4.5.3. Strategies and methods to reach clinical judgements and decisions

There were several types of methods identified that were used to make a diagnosis and/or to make a decision to manage a patient. This was a reflection of the complexity of clinical reasoning and the repertoire of multiple cognitive processes used by doctors. It was unclear how doctors selected which method was most appropriate to arrive at a diagnosis or decision. For instance, some management decisions were based on a doctor's preference or an ingrained habit. At other times, doctors would deliberate about the pros and cons of multiple options before making a choice. According to the Cognitive Continuum Theory (Hammond, Hamm, Grassia, and Pearson 1987) the decision maker uses different styles of reasoning depending on the nature of the decision task. A distinction is made between a well structured and an ill structured decision task. This theory proposes that well structured decision tasks are complex and require time and effort to resolve. This decision context primes a person to use a conscious systematic processing method. Whereas an ill structured task is less complex and can be solved quickly, such tasks are said to invoke heuristic processing where the decision is reached without much conscious awareness. On the surface, this may be a suitable explanation for the present findings, but it is unclear as evidence exists that contradicts the cognitive continuum theory (Kulatunga-Moruzi, Brooks, and Norman 2004; Pretz 2008; Ark et al. 2007).

4.5.4. Perceptions about the development of clinical reasoning competency

Participants unanimously held a traditional view of expertise, that clinical reasoning improved over time with more clinical experience (Ericsson 2007). It was learned through imitation of colleagues, repeated practice with real patients and knowledge of accepted procedures at the workplace. In other words, doctors learned how to make clinical judgements and decisions through the hidden medical curriculum rather than the formal curriculum. Participants also felt biomedical knowledge was necessary, but to what extent was unclear. In the literature, there are differing opinions about the precise role biomedical knowledge plays in clinical reasoning (Woods 2007). In this study junior doctors were more mechanical in their thinking and rule bound in their practice, compared to senior doctors. Similar findings were reported about other junior doctors and nurses (Benner 1982; Jensen et al. 1990; Shanteau 1992; Greenhalgh 2002). Similarities between junior and senior doctors were noticed as well. Like senior doctors, junior doctors and medical students referred to their use of pattern recognition as the default method of reaching a diagnosis. This suggested that novices had developed enough clinical experience to recognise, at least some conditions, based on their collection of signs and symptoms. Evidence shows that encountering similar conditions repeatedly is a critical component in the development of expert diagnostic reasoning (Norman, Rosenthal, Brooks, Allen, and Muzzin 1989; Brooks et al. 1991; Hatala, Norman, and Brooks 2003). Some authors have concluded that the recognition of patterns is how routine diagnosis is made by all doctors (Elstein & Schwartz 2002; Norman 2009). Another similarity between participants was that most, including senior doctors, did not believe that expertise was an end state. Senior doctors did not perceive themselves to be expert decision makers nor did junior doctors believe they would become expert decision makers. Rather, they felt it was important for them to recognise their limitations and remain as continual learners. This view resembled one type of expert identity that has been identified, known as the adaptive expert (Mylopoulos & Woods 2009).

4.5.5. Implications of findings to medical education and practice

This study has highlighted several implications for medical education and practice. The doctors drew upon a large range of cues when thinking about the cause of a condition or how to treat a patient. Some of these cues were helpful – such as when it related to disease probability – other cues were potentially misleading – such as the *strangeness* of a patient's personality. Only a few doctors demonstrated they were aware of how some factors could lead them to make biased choices. It would be difficult to teach doctors about the cues that should definitely be considered and those that should definitely be ignored. A patient's appearance was described as an important source of information about their health, but this could also be a distraction if it led to a biased choice. Instead, doctors should be made aware that they selectively process information, that is, they do not consider all of the available information. The information that they do consider is not always relevant to making an accurate diagnosis and/or selecting the best treatment option.

There was a consensus that a good judgement or decision was reasoned well. This definition presents practical problems for measuring a good reasoning process as there is no common understanding of good and bad thinking (Bekker 2010). In the systematic review study, it was found that few studies measured the process of clinical reasoning. Of those that did, good reasoning about clinical problems were indicated by avoiding biased judgements when making diagnoses (Hershberger et al. 1995; Round 1999; Abraham et al. 2004) or following a normative guideline of good prescribing skills (DeVries et al. 1995; Hassan et al. 2000; Akici et al. 2003). Others have offered explicit criteria of good reasoning (Pauker & Pauker 1999).

It is established that doctors draw on a repertoire of different methods when thinking about clinical problems. What is not established is whether doctors can be helped to identify the optimum mode of thinking for a particular decision problem. It has been suggested that doctors should be aware that they use more than one style of reasoning. Further evidence suggests greater diagnostic success was associated with explicit instruction to use a combined reasoning approach of both non-analytic and analytic strategies, compared to when no instruction was given (Ark et al. 2006; Ark et al. 2007; Eva et al. 2007). It is now being recognised that different types of diagnostic strategies should not be thought of as mutually exclusive and that doctors should not limit themselves to one mode of thinking (Eva et al. 2007). Most participants were receptive to receiving formal training about decision making, but stressed that it must not be so theoretical that they could not recognise the relevance to their work. Some preferred it to be integrated into clinical practice as a work-based learning activity. It was interesting that having teaching responsibilities was perceived to be a benefit to doctors' clinical reasoning. Whether it actually helps them make better judgements and decisions than doctors who do not teach others, is a topic for future research.

4.5.6. Critique of study

The study's strengths include the rigorous methods used and in-depth data generated to further understanding of clinical judgement and decision making. The sampling frame ensured that a range of doctors' views about their experiences of making judgements and decisions, were represented. Steps were taken to develop a suitable interview schedule in accordance with good practice, to encourage participants to speak freely, in depth, and without bias (Robson 2002; Ritchie & Lewis 2003). The data and themes generated are thought to be robust as the materials and procedure were piloted, and the coding frame was developed according to an established qualitative methodology (Ritchie & Lewis 2003). That the present findings are in line with similarities identified in other literature further supports the validity of the data.

The study may be subject to the limitations associated with qualitative methods. Findings from the study may not be applicable beyond the study sample for two main reasons. First, this was not an observational study. This means the data reflects doctors' perceptions of how they make judgements and decisions and not necessarily how they are actually made in practice. Second, even though effort was made to interview doctors from a range of specialties and levels of experience, it is unclear whether these views and experiences about making clinical judgements and decisions, are common to all doctors. There is also a chance that the themes generated were influenced by the author's prior knowledge about the topic area. However, effort was made to code the data according to the words and meanings of the participants. This provides some confidence that the themes are an accurate reflection of participants' views and experiences about clinical decision making.

4.6. Summary

This study provided deeper insight into the complexity of how doctors experience making clinical judgements and decisions. The sample had limited awareness about the processes doctors employ to solve clinical problems and there was no process described as ideal. There was some awareness that a diverse array of cues influenced their thought processes, but little explicit awareness about the occurrence of biased thinking. A general competency in clinical reasoning was believed to principally develop through years of repeated clinical experience, but all except one participant supported the introduction of formal training to improve clinical reasoning into medical education.

5.

Evaluating the feasibility of an online decision literacy intervention: a quasi-experimental study

This chapter describes a quasi-experimental study to evaluate the feasibility of delivering an online intervention to enhance doctors' decision literacy. Participants were shown a short tutorial and worked through a series of tasks. The tutorial provided basic information about how people use a heuristic and/or systematic method to make judgements and decisions. Clinical problem solving exercises and multiple choice questions were used to evaluate the intervention based on the following outcomes; knowledge of decision sciences concepts, relevance of the tutorial material and clinical judgement.

5.1. Background

From the synthesis of evidence in Chapter 2 the following aspects seemed suggestive of an effective clinical reasoning course for doctors: an explicit theoretical basis, appropriate assessment methods, useful feedback and integration into medical education. An understanding of how people in general and doctors specifically, make judgements and decisions is required, as discussed in Section 1.5. Knowledge about the delivery of medical education is also necessary to appropriately design and evaluate an intervention. These issues are discussed further below.

The findings of the systematic review and questionnaire studies implied that educators focus on training doctors to apply normative decision theory to clinical problems. This included teaching how to calculate disease probabilities using Bayes theorem (e.g. (Kurzenhauser & Hoffrage 2002; Grant et al. 2006) and how to apply decision analysis to find the optimal treatment choices (e.g. Lee et al. 2007). There has been little focus on enhancing doctors' understanding of their own reasoning processes, in spite of the encouraging findings (Wolf et al. 1988; Hershberger et al. 1995; Round 1999; Abraham et al. 2004). This type of intervention could be informed by the decision sciences, including content about information processing (Section 1.5.1) dual processing models of thinking (Section 1.5.4) and the heuristics and bias literature (Section 1.6).

In terms of training, the information processing approach and dual processing models provide a useful theoretical basis for understanding how to intervene in doctors' clinical reasoning (Dowie & Elstein 1988; Croskerry 2009b). They explain how people make lay and professional judgements and decisions as well as how they can be improved. There is a hypothesis that if doctors are educated about their use of heuristics, they may be able to avoid the influence of bias (Arnoult & Anderson 1988; Croskerry 2000; Croskerry 2003). From the systematic review, there is evidence that raising doctors' awareness of heuristics and bias can improve aspects of their clinical reasoning on hypothetical scenarios (Wolf et al. 1988; Gruppen et al. 1994; Round 1999; Hershberger et al. 1995; Abraham et al. 2004). A potentially encouraging finding was reported by Bornstein et al. (1999). They found that doctors were not influenced by a sunk-cost effect in medical situations, but were in non-medical situations. Specifically, their treatment decisions were not biased by the amount of time and money already invested in treating a patient, but these factors did bias their non-medical decisions (Bornstein et al. 1999). However there is little evidence to suggest that gaining knowledge about heuristics and bias would improve actual judgements and decisions, in practice. In the non-clinical context, awareness training about heuristics was not enough to improve their judgements in practice (Fischhoff 1975; Welsh et al. 2007).

As well as theory, the design and evaluation of an educational intervention should be informed by current trends in medical education. There has been a shift from traditional lecture based teaching to problem based, student self-directed learning (Gwee 2003). Web-based learning, has become accepted within the medical education community and is widely used as a supplement to traditional teaching (Ruiz et al. 2006). In the previous systematic review study, 14 studies delivered their intervention electronically and six of these were web based tutorials using clinical vignettes (Warner et al. 1974; Murray et al. 1977; Lincoln et al. 1991; Kumta et al. 2003; Servais et al. 2006; Hedrick & Young 2008). Five of these web based interventions reported that doctors performed better in the training groups compared to the control groups. Further, a review of 76 studies from the medical and dental literature found that web based learning was as good as, but not better than traditional teaching (Chumley-Jones, Dobbie, and Alford 2002). Web based learning can be used to encourage doctors to take greater responsibility for their own learning and engage with continuing professional development activities. To encourage problem based and active learning, assessment methods should target not only knowledge components such as recall tests but methods that demonstrate what trainees will do when faced with real clinical problems (Miller 1990).

An important educational strategy is the integration of teaching throughout medical education. It is necessary to consider where and how clinical reasoning training could be included in medical education, as it is likely to be a key factor in the delivery of an effective educational programme (Harden 2000). The systematic review and questionnaire studies found that most courses were not integrated into medical curricula. Those that had been integrated were no more effective than the one-off studies, suggesting effective integration is a complex issue that requires careful planning. General guidance has been offered on how to integrate educational programmes in medicine (Harden 2000) but little is known about where to place training in clinical reasoning into medical education.

Frameworks have been designed to guide the evaluation of educational training programmes. According to Kirkpatrick's model and Miller's framework of clinical assessment, training should be evaluated according to a hierarchical sequence of four levels (Miller 1990; Kirkpatrick 1998). Kirkpatrick suggests that an intervention should at least assess participants' reaction to the training, followed by their learning. At the higher complex levels, training should be assessed in terms of behavioural changes and results in the workplace (Kirkpatrick 1998). Miller (1990) describes four stages of competencies a doctor should demonstrate in clinical assessment. At the lowest level a doctor is expected to have gained knowledge, followed by knowing how to use that knowledge. Assessment at the higher levels becomes more difficult, when doctors should actually demonstrate they can correctly use the knowledge gained, followed by applying it appropriately in clinical practice (Miller 1990).

5.2. Aims and objectives

There is no study that evaluates an educational intervention to explicitly teach doctors about the processes they use to make judgements and decisions. This study aimed to enhance doctors' awareness of how they make personal and clinical judgements. In other words, the intervention aimed to enhance doctors' decision literacy. It was informed by the decision sciences and the medical education literature. The study also drew upon findings from the previous empirical studies in Chapters 2, 3 and 4. It explored whether it was useful to teach doctors about information processing and factors that influence their judgements and choices. This insight may help doctors reflect on their clinical judgements and decisions. However, it may not be enough to change practice. The study sought to answer the following research questions:

- Does an intervention about basic decision science improve doctors' knowledge of how they make judgements and decisions?
- Can the intervention be delivered in a way that is relevant to doctors?
- Does the intervention improve doctors' clinical judgements about diagnosis?

5.3. Method

5.3.1. Context

The intervention was designed online in the format of web-based learning for the following reasons. First, doctors are familiar with participating in online courses, assessments and completing e-portfolios to monitor their professional progress. Second, it was a convenient and accessible way for doctors, who had busy schedules, to participate in the study. The only requirement was access to a computer with Internet access. Third, participants could work through the study at their own pace without the presence of a researcher. This was expected to reduce the likelihood of biased responses that occur from researcher effects (Robson 2002). Fourth, it linked well with doctors' requirement of continuing their professional development (CPD). Doctors are also accustomed with a case-based learning format using clinical vignettes and multiple choice questions (MCQs) to assess their learning. For this reason, clinical vignettes and MCQs were used to evaluate the online intervention.

5.3.2. Ethical approval

The study received ethical approval in 2010 from the University's Medicine and Dentistry Educational Research Ethics Committee (EdREC) (Appendix 8.12).

5.3.3. Design

The study used a quasi-experimental pre-test/post-test design. Experiments are the most appropriate design to answer questions about the effectiveness of interventions (Gribbons & Herman 1997).

The ideal type of experiment is the randomised control trial (RCT) also known as a true experiment (Coolican 1999; Robson 2002). Its key features are a control group and random allocation of participants to each group. In comparison to other study designs, the true experiment enables a researcher to maintain more control over extraneous variables that may lead to a misinterpretation of results. The advantage of a high level of control is that any difference found between control and experimental group can be reasonably attributed to the manipulation of the experiment (Kantowitz, Roediger, and Elmes 2005). However it is not always feasible or appropriate to carry out a true experiment, especially in education contexts. A suitable alternative is a quasi-experimental design (Cook & Campbell 1979). A quasi-experimental design does not require random allocation of participants. In the present study the participants were only contacted once and it was anticipated that the sample size would be too small for a full randomised controlled trial. A quasi-experimental design is also well suited for applied research where some flexibility is required in the design (Robson 2002). For this study, a quasi-experimental design was seen as a compromise between flexibility whilst still maintaining an experimental approach. Questionnaire methods were used to evaluate the impact of the intervention. Questionnaires offered a simple way of collecting a large amount of standardised data. They can also be completed online which was a necessity in this study.

A pre-test/post-test design was used with the same group of people. The advantage of this design is that it eliminates the potential for individual differences to confound the results (Coolican 1999). When there are multiple groups of different people, individual differences in participants may be unevenly distributed. A biased interpretation of results is possible if differences in results are found. The disadvantage of a quasi-experimental pre-test post test design is that there is less control over extraneous variables than in true experiments. This means that it is not possible to attribute with certainty any change in results to the intervention, rather it can only be concluded that the variables are related (Kantowitz et al. 2005). For instance, there is a possibility that results may improve after the intervention due to completing the same task twice, through practice effects (Robson 2002). Nevertheless, the quasi-experiment allows at least some knowledge to be gained about the effectiveness of an intervention when a true experiment is not feasible (Cook & Campbell 1979).

5.3.4. Sample

Medical doctors who were either postgraduate students or staff members at the University of Leeds were invited to take part in the study. To encourage participation, the study was aimed at doctors who were also students. As the context of the study was based on doctors' continuing professional development, it was assumed postgraduates had an interest in continuing medical education. Further, they had experience of making real clinical judgements and decisions. Undergraduate medical students were not invited to take part as they did not have sufficient clinical experience to complete the tasks. The sample size estimate was based on commonly used rules of thumb used in statistics. Some authors suggest that to compare differences between groups, there should be a minimum of 7 people in each group in order to maintain sufficient power in the study (Wilson VanVoorhis & Morgan 2007). Others suggest 5 participants per variable are sufficient when estimating a minimum sample size (Tabachnick & Fidell 1996). In this study there were two independent variables that were manipulated: (i) web based tutorial and (ii) type of clinical vignette. There were three dependent variables: (i) knowledge of decision sciences concepts, (ii) relevance of the tutorial context, and (iii) clinical judgement. Based on the rule above of five participants for six variables in total, it was calculated that the study sample should be a minimum of 25 participants.

5.3.5. Materials

5.3.5.1. Study information and consent

An information sheet that described details about the study was designed to recruit participants (Appendix 8.13). Information sheets were distributed electronically to potential participants and hard copies were handed out during teaching sessions by the author. The content of the information included an introduction of the author, description of those invited to participate, purpose of the study, web link of the study and what the study involved, information about a prize draw, what happens to the data and contact details of the author. When volunteers accessed the web link they were presented with a summary of the information again as an introduction to the study. Before beginning the tasks volunteers were required to confirm they had understood the information and give their consent to participate.

5.3.5.2. Overview of the intervention

The online intervention consisted of three main aspects. First, there were problem solving exercises based on clinical vignettes. These were used to assess doctors' judgements about diagnosis in a way that was familiar and relevant to their continual professional development. Secondly, there was a tutorial to improve doctors' understanding about how they make personal and clinical judgements and decisions. Information about basic decision science was provided using real examples of lay judgements and examples that were relevant to their clinical practice. Third, there was an evaluation of participants' knowledge of decision concepts and the perceived relevance of the tutorial material.

The intervention was implemented using the Bristol Online Survey (BOS) application. This service allows questionnaires to be developed and deployed online. There were three reasons for why this application offered the most practical method to develop the experiment online. First, the University of Leeds has an annual subscription to BOS which enables staff and students to obtain a password protected account free of charge. This meant there was a good support network within the University if any problems arose with the use of BOS. Second, the data are stored on a secure server at the University of Bristol and, thirdly, results can be exported into SPSS. It was not possible to design the tutorial using the BOS application, so this was designed as a series of presentation slides using Microsoft PowerPoint 2003. This PowerPoint presentation was uploaded onto the author's personal web space at the University of Leeds, and the URL link was embedded into the BOS application. A disadvantage was that the tutorial was less integrated into the experiment than the other tasks. Alternative online applications were considered such as SurveyMonkey and Questionmark, but these did not meet the needs of the study design nor were they supported by the University of Leeds.

5.3.5.3. Designing the tutorial

The purpose of the tutorial was to enhance doctors' understanding of how they make judgements and decisions in non-clinical and clinical contexts. The content of the tutorial was based on evidence and theories from the decision sciences. The tutorial consisted of fourteen PowerPoint slides (Appendix 8.14). It began with introducing the purpose of the tutorial and why it was relevant to doctors. Learning objectives were given so that participants were aware of the knowledge they should have gained on completion of the tutorial. The concept of information processing and limited capacity to process information was explained. It then moved onto the main content of the tutorial regarding a heuristic and systematic method of processing information. In order to facilitate participants' understanding of the details about heuristic and systematic processing methods, numerous examples from a non-clinical and clinical context were given. This illustrated to participants how they use these methods to make judgements and decisions in their personal lives, as well as in their clinical practice. Images were also included that depicted examples of the use of these two methods to reinforce understanding and add interest to the tutorial. A table was provided that compared the main features of a heuristic and systematic method so participants could see the differences. The tutorial ended with a take home message and a summary of the key concepts that had been described previously.

5.3.5.4. Outcomes

There were three outcomes that were measured to evaluate the intervention:

- knowledge of decision sciences concepts;
- relevance of the tutorial;
- clinical judgement about diagnosis.

In accordance with Kirkpatrick's model, the outcome perceived relevance was a measure of participants' reaction (Kirkpatrick 1998). In accordance with Miller's framework, the outcome knowledge of decision making concepts measured participants' knowledge gain from the tutorial (Miller 1990). Clinical judgement measured whether participants could demonstrate correct use of the knowledge gained.

5.3.5.5. Measuring knowledge of decision sciences concepts

A questionnaire was developed to measure participants' knowledge of decision making concepts using a questionnaire format (Appendix 8.15). This measure included a series of 10 multiple choice questions about the decision making concepts presented and was administered after the tutorial. Participants were asked to identify the name of the concept described in each statement. They had to select one of the following responses each time: systematic strategy, information processing, smart heuristics, heuristic and systematic strategy, limited capacity, or heuristic strategy. Responses were in the format of multiple choices as doctors are usually familiar with this as a method of assessment. Also, the ability to recognise the correct answer was a valid indicator that the tutorial content was understood. Concepts from the decision sciences presented in the tutorial were complicated and perhaps new information for most people. It would have been unrealistic to expect participants to memorise all of the terms introduced in a five minute tutorial. Therefore they were not asked to freely recall any of the tutorial information. Participants were given the opportunity to show they had understood the tutorial content through recognising the correct description of each key concept.

5.3.5.6. Measuring relevance of tutorial

A questionnaire was developed to measure the extent to which participants found the tutorial content useful and applicable to their clinical work. It consisted of 13 statements that participants were asked to rate their level of agreement with (Appendix 8.16). This measure was also administered after the tutorial. The following items on the questionnaire assessed participants' overall perception of the benefit tutorial;

The tutorial was interesting.

I do not feel I have benefitted from the tutorial.

The tutorial taught me nothing new.

I think some formal teaching on how people make decisions should be included in medical training.

The following items assessed whether participants understood the relevance of the tutorial information to their own clinical practice;

It helped me to think about my own decision making.

I have a better understanding of how people make decisions.

I understood the relevance of the tutorial to me as a doctor.

I have a better understanding of how others can affect my judgements and choices.

I have a better understanding of how my personal beliefs and experiences can affect my judgements and choices.

Further, participants were asked questions around how the information may actually affect their clinical practice with the following:

I can identify examples of when people use heuristic and systematic strategies to make personal decisions.

I can identify examples of when doctors use heuristic and systematic strategies to make clinical decisions.

I feel more cautious about making clinical decisions.

I feel more confused about making clinical decisions.

5.3.5.7. Measuring clinical judgement

To measure the impact on clinical judgement, a set of four problem-solving tasks were developed using a questionnaire method (Appendix 8.17). The problem solving tasks consisted of two components, clinical vignettes and a series of associated multiple choice questions for participants to answer. This measure was administered before and after the tutorial to evaluate whether there was any change in participants' clinical judgements about their diagnosis. The details of how the components of this measure were developed are discussed below.

5.3.5.8. Developing clinical vignettes

Since the inception of problem based learning, clinical vignettes have been a commonly used teaching tool for doctors. They are relevant, practical examples of real situations that occur in the clinical context. There is evidence that vignettes are a valid way of measuring doctors' ability to diagnose and treat specific medical conditions (Peabody, Luck, Glassman, Dresselhaus, and Lee 2000; Peabody et al. 2004). Clinical vignettes were also used in many of the interventions reviewed in Chapter 2. They were used for a variety of purposes such as to illustrate cognitive biases in clinical reasoning or to practice making good prescribing decisions (DeVries et al. 1995; Round 1999). Others used them to train doctors in improving their clinical diagnostic judgements (Wolf et al. 1988; Wigton et al. 1990; Friedman et al. 1999).

The content of the clinical vignettes were informed by anecdotal evidence in a recent bestselling book titled *How Doctors Think* (Groopman 2007) as well as actual scenarios taken from the interview study (Chapter 4). Four vignettes were developed that were based on actual scenarios. Two vignettes were based on scenarios taken from the book *How Doctors Think* (Groopman 2007). Groopman is a senior doctor and academic who has published over 150 articles in peer reviewed journals. He is also involved in 'popular' medical writing for various newspapers such as *The New Yorker*, the *New York Times* and the *Washington Post.* The book *How Doctors Think* was particularly relevant to this thesis as Groopman has studied the heuristics and bias literature extensively. He uses anecdotes from doctors and patients to illustrate the times when doctors made good and poor judgements and explores the reasoning processes used.

Vignettes 1 and 2 were developed from clinical scenarios narrated by two medical students during the interview study, reported in Chapter 4. These were taken from the transcripts of interviews 12 and 13. Vignettes 3 and 4 were based on two scenarios narrated in Groopman (2007). They described real situations when two experienced doctors failed to diagnose their patients in an accurate and timely manner, which almost led to fatal consequences. The clinical vignettes were as follows:

> **Scenario 1:** A 19 year old man comes to the Accident and Emergency department (A&E) late at night complaining he is struggling to breathe. He is a student and this has been his first week at university. You take his history and find out he had mild asthma as a child, but feels he has grown out of it with age. You examine him and he sounds wheezy and is taking quick short breaths. He states he has no chest pain and has had no recent illness. (Int. 13)

> **Scenario 2:** A mother brings her 1 year old baby into the GP surgery. She explains her baby has developed a fever and bad cough yesterday which has worsened during the night. As the baby cries it starts coughing and you notice it sounds like a bark. You take the baby's temperature and it is slightly raised. The breathing is noisy, but the baby does not appear to be struggling for breath. There is no rash on the body and throat and ears look normal. (Interview 12)

> **Scenario 3:** A young man is brought into Accident and Emergency in the early hours of one morning. The police had found him slumped on the steps of a public library. He is unshaven, his clothes are dirty and is not fully conscious. He is unwilling to arouse himself. He seems confused and cannot respond with any clarity to the nurse's questions. (Groopman, 2007)

> **Scenario 4:** A 41 year old man comes into Accident and Emergency complaining of severe chest pain. He explains he was hiking in the

woods when a pain in his chest stopped him in his tracks. He works as a forest ranger so due to his active lifestyle has a lean figure and is very fit. Over the past few days he has experienced growing discomfort in his chest, even when resting. As a forest ranger he is used to muscle aches but thinks this is different. He has no history of heart problems. (Groopman, 2007)

All of the vignettes were based on real situations that had taken place between a doctor and patient. It was preferred that the scenarios were authentic rather than fictitious in order to find out how doctors might judge these situations in the clinical setting. The study was aimed at doctors working in any specialty so it was important that the scenarios did not represent presentations that were specialty specific. The scenarios were selected because they were based on general medicine presentations. It was anticipated that participants would be familiar with presentations that resembled an asthma attack, croup, diabetes and angina. This was to ensure that all participants would at least be familiar with the presentations so they could respond to questions in a considered rather than random manner. If the vignettes described conditions that were particularly rare or specific to one specialty, then some participants would be unfairly disadvantaged if they were unaware of that presentation. If this was the case, people may guess answers to questions without much thought, potentially introducing bias into the results.

The vignettes were written in a way to reveal only a limited amount of information about the patient's condition. This was to assess the initial judgements participants made about each scenario. The ambiguity of the vignettes was manipulated for the purpose of distinguishing whether the tutorial had an overall positive rather than negative effect. It was important to establish whether the intervention had a detrimental effect on the correct judgements people made before the tutorial. This was not a desired outcome and would mean the intervention would be ineffective and inappropriate to include in medical education. Scenarios 3 and 4 were more ambiguous than one and two. To achieve this, an understanding of how doctors make diagnostic judgements was necessary. Evidence suggests that diagnosis is usually a direct and automatic process based on pattern recognition (Kulikowski 1970; Patel & Groen 1986; Norman 2005). That is, doctors are able to recognise the signs and symptoms of many medical conditions, without much conscious thought. In scenarios 1 and 2, it was expected that the correct diagnosis would be easily recognised by most participants. The signs and symptoms given were typically characteristic of a patient that had an asthma attack in scenario 1, and a baby that had croup in scenario 2. Whereas in scenarios 3 and 4, it was expected that the correct diagnoses would be less recognisable to participants. The signs and symptoms represented atypical presentations of a patient with diabetes in scenario 3 and angina in scenario 4. Further, the signs and symptoms given also resembled those of other conditions i.e. the presentations were ambiguous. Specifically, in scenario 3 it was expected that some participants would misjudge a young, semi-conscious man to be drunk rather than correctly diagnose him as diabetic. Information such as his unkempt appearance and being found slumped on the steps of a public library, were expected to bias their judgement towards an incorrect diagnosis. In scenario 4, it was expected that some participants would misjudge the cause of chest pain in a young forest ranger to be muscle strain rather than angina. Information such as his active lifestyle and seemingly high level of fitness, were expected to bias their judgement towards an incorrect diagnosis. These assumptions were based on the actual misjudgements that doctors made in practice as narrated in Groopman (2007).

5.3.5.9. Developing multiple choice questions

After each clinical vignette, participants were asked to respond to the same set of four multiple choice questions. They were asked the following questions about their diagnostic hypotheses for each scenario.

1. Given these findings only, what diagnosis would you make?
2. How certain are you that this is the correct diagnosis?
3. Would you actively wish to exclude or diagnose another condition? If so, which condition(s)?

4. Rate how much influence the following aspects of the scenario had on your chosen diagnosis?

The purpose of questions 1 and 2 was to understand participants' first impressions of the scenarios and also whether their choice of diagnosis changed after the tutorial compared to before. The purpose of questions 3 and 4 was to understand how participants reasoned about the diagnosis they had chosen.

5.3.6. Evaluating quality of materials

It was necessary to evaluate the quality of the materials. Steps were taken to assess the validity of the measures and the suitability of the online format of the intervention. This included discussions with supervisors and a review of the materials by a group of experts. The final stage was a pilot test of the materials and study procedure by a group of doctors that did not participate in the main study. Further details of these stages are discussed below.

The items that measured perceived relevance were counter-balanced. Some were positively worded such as 'I have a better understanding of how people make judgements and decisions', and others were negatively worded such as 'I do not feel I have benefitted from the tutorial'. This was to encourage participants to think about their responses rather than answer in an automatic manner. People have a tendency to agree with items when asked to rate their level of agreement, known as acquiescence bias. Counterbalancing the way the items are worded or scored is a way of reducing the effect of this bias (Robson 2002).

To ensure that clinical judgements were accurately operationalised by the problem solving tasks, they were reviewed by supervisors and three experienced GPs that worked at the Leeds Institute of Health Sciences. GPs were appropriate to consult because the clinical scenarios were based on presentations that would be seen in general medicine. They were asked to provide feedback about the appropriateness of the types of clinical presentations i.e. were they familiar to doctors irrespective of their specialty and seniority level;

was the length of the scenario suitable; and was the language comprehensible? All agreed that the scenarios were authentic and the presentations would be familiar to all types of doctors. This provided some confirmation of face validity.

Their feedback was also sought on the wording of multiple choice questions and the different response options. A few ambiguities were identified and amended. It was suggested the first question regarding diagnosis, should be changed from *What do you think is the most likely diagnosis?* to *Given these findings only, what diagnosis would you make?*. The former was judged to be potentially misleading and may have biased participant responses. An example given was that in scenario 4, alcohol excess would arguably be the most likely cause because it is statistically a more frequent presentation in A&E than a patient on the verge of a diabetic coma. To avoid this bias, the question was re-worded in more neutral terms. Another question was changed from *Would you want to obtain further investigations for another condition? If so which condition(s)?* to *Would you actively wish to exclude or diagnose another condition? If so, which condition(s)?*. This change was made because not all of the other possible conditions listed required actual investigations, which prompted clarification in the wording of the question. The relevant changes were made and sent back to the group of experts for further consideration. The amended version was judged by the experts to be a suitable test of doctors' judgement when giving diagnoses. This provided evidence that the scenarios and multiple choice questions were an accurate representation of real clinical judgements made in the clinical context i.e. expert opinion suggested the construct validity of the questionnaire. Further, the difference in ambiguity of the scenarios was judged to be an appropriate way of distinguishing any change in judgements after the tutorial, again a sign of construct validity.

The final step in establishing validity in the intervention was to pilot test the materials and study procedure. Using the BOS application, the pilot experiment was designed. A group of doctors that worked in the Leeds Institute of Health Sciences were invited to participate. These were not the same people who helped design the clinical vignettes. Five doctors participated in the pilot experiment. They were asked to provide feedback on the length of time it took to complete the study, ease of comprehension of instructions and completion of tasks and any problems that occurred. The feedback received was mostly positive. Participants reported the study was interesting and insightful. It took people between 15 and 20 minutes to complete the study and this was reasonable given the number of tasks to complete. The instructions were described as clear and overall the experiment was easy to complete. The results were examined and participants answered in a similar manner to each other before and after the intervention, suggesting there was some reliability in the clinical vignettes. Patterns emerged that were anticipated. Participants did select the correct diagnosis before and after on the unambiguous scenarios and were fairly certain both times. This suggests there was some reliability in scenarios 1 and 2 and that the intervention was not associated with a detrimental effect on their correct judgements. It was expected that some people would make biased judgements about the ambiguous scenarios and this pattern also emerged.

5.3.7. Procedure

Six course tutors were approached by the author for permission to advertise the study amongst their students. Those that agreed sent an e-mail to their students to advertise the study on behalf of the author. The study information sheet was distributed amongst students of the following courses: Postgraduate Certificate in Health Research, Masters in Psychiatry, Masters in Public Health, Masters in Public Health for International Students, Masters in Child Health and Masters in Clinical Education. Additionally, an e-mail was sent to staff members of the Faculty of Medicine and Health addressed to those who were also medical doctors.

The study was completed in four parts. Participants accessed the web link that had been provided by the author in order to access the study online. Part one required that people read the welcome page that reiterated information about the study. They were told the study was expected to last approximately 20 minutes. Participants were required to confirm that they had understood the information and that they agreed to take part in the study. They were given information about a prize draw and were instructed to provide their email address so they could be contacted if they had won a prize. The first task began with the clinical problem solving exercises. Participants were required to work through each scenario and all of the associated questions, described in the materials section. This task was expected to take 5 minutes to complete.

In part two of the study, participants were shown the tutorial. They were instructed to click on the web link to access the tutorial and then close it once they had read the information presented. It was recommended that participants spend at least 5 minutes reading the information given in the tutorial, to get the most benefit. In part three of the study, participants were asked to complete the same clinical problem solving exercises that were given to them before the tutorial. This was to determine whether students thought differently about the same scenarios after the tutorial compared to before. Participants were told that they could choose the same responses as they did previously, change some of them, or change all of them. The final part of the study required participants to complete an evaluation questionnaire. This included multiple choice questions to assess their understanding of the tutorial content and a section that asked them to rate how useful they found the tutorial. This task was also expected to take approximately 5 minutes. After this was completed, participants were informed that they had reached the end of the study and that their answers had been submitted.

5.4. Analysis

Results were analysed using SPSS for windows version 18.0. All participants completed the study fully, there was no missing data. Results were described in terms of descriptive and inferential statistics.

5.4.1. Analysis of knowledge and relevance

Knowledge of decision sciences concepts and perceived clinical relevance were measured on the second questionnaire, after the tutorial. Participants were measured via multiple choice questions about their knowledge of decision making concepts. The total number of correct responses out of 10 was entered into SPSS for each participant. As the data was continuous in nature, the mean number of correct responses was calculated along with the standard deviation. Frequencies were also calculated to show how many participants scored a particular total number on the multiple choice test, for example how many people scored 5, 6, 7 etc. Participants also rated their level of agreement with a series of statements. This indicated their views about the clinical relevance of the tutorial information. This task obtained ordinal data. The frequency of people that agreed, disagreed or were neutral about each statement, were calculated.

5.4.2. Analysis of clinical judgement

Participants were asked questions about their diagnostic judgements on four clinical vignettes. There were four questions to answer for each vignette, therefore 16 questions in total. Participants' responses for each question were coded as numeric variables in SPSS and crosstabs were computed to describe the results. Crosstabs were particularly useful as they display frequency data in a contingency table to express the association between two variables (Bryman & Cramer 2008). This was an accessible way of displaying the responses participants chose for each question on each scenario before and after the tutorial. Not only did crosstabs provide information about how many people selected a particular response option, but also details of the changes in people's responses after the tutorial.

In order to determine if there were any significant differences in the way participants responded to questions before and after the intervention, statistical analyses were conducted. The data were categorical and ordinal in nature. Categorical variables were associated with questions that asked people to select a diagnosis and other conditions they may wish to exclude or diagnose. The response options to these questions were in the form of categories that had no intrinsic order to them (Bryman & Cramer 2008), for example asthma attack, panic attack, pneumonia and heart attack. In order to determine whether responses before and after the intervention differed on these categorical variables, the McNemar's chi-square test was conducted. The McNemar chi-squared test is a non-parametric test used to analyse categorical data (Bryman & Cramer 2008). It is also used when the same individuals have been measured twice on the same variables, as was the case in this study.

Ordinal variables were associated with questions that asked participants to rate the certainty level they had towards their chosen diagnosis and the influence each piece of scenario information had on their diagnosis. The response options to these questions were in the form of categories that could be rank ordered, but in which the differences between categories were not equal (Bryman & Cramer 2008). For instance, not at all certain, somewhat certain, fairly certain and absolutely certain. The Wilcoxon signed rank test was conducted to determine whether responses before and after the intervention differed on these ordinal variables variables. The Wilcoxon test is an alternative non-parametric test to the repeated measures t-test. For a t-test to be used appropriately, the data should be (i) continuous i.e. numbers can be compared as multiples because differences between numbers are identical and (ii) have a normal distribution. This study did not meet those conditions. The Wilcoxon test was more appropriate as it is designed to calculate differences when data is ordered in ranks and does not make assumptions about the distribution of the data (Bryman & Cramer 2008).

5.5. Results

The study was complete by 48 doctors who were either staff members or postgraduate students at the University of Leeds. Half of the sample provided demographic data about themselves. There were 14 females and 10 males between the ages of 29 and 50. The sample of doctors varied in level of seniority and clinical specialty. Most participants worked at registrar level, and ranged from those who were newly qualified to having seven years registrar experience. Only a few people were consultants. Clinical specialty varied widely from Psychiatry, Surgery, Oncology, Cardiology, Paediatrics, Rheumatology, Renal Medicine, General Practice and a few were in academic positions. The results were summarised under three main headings that reflected the outcomes measured. These were knowledge of decision making concepts, clinical relevance of the tutorial and impact on clinical judgements.

5.5.1. Knowledge of decision making concepts

Most participants did well on the multiple choice test. All participants achieved at least 50% correct answers, and 79% scored 8 or above. The mean score was 8.33 (standard deviation = 1.34). One third of the sample scored 9 and this was the most frequent score. The lowest score was 5 and the highest was 10.

5.5.2. Relevance of tutorial

Over half of the sample found the tutorial interesting and felt that that they had benefitted from it and were taught something new (Table 27). The majority agreed that some formal teaching on how people make judgements and decisions should be included in medical training. In terms of understanding the content of the tutorial, just over half agreed that it helped them to think about their own clinical judgements and decisions and over 70% of people understood the relevance of the tutorial to doctors. However, only a small minority of people agreed that they had gained better understanding of how people make judgements and decisions. Most people were ambivalent about this and approximately one third disagreed that they had gained better understanding about how people make judgements and decisions. Overall people seemed unclear about the factors that can affect their judgements and choices.

Statement	Disagree	Neutral	Agree
The tutorial was interesting	19%	23%	58%
It helped me to think about my own judgement and decision making	19%	27%	54%
I have a better understanding of how people make judgements and decisions	38%	56%	6%
I understood the relevance of the tutorial to me as a doctor	14%	13%	73%
I can identify examples of when people use heuristic and systematic strategies to make personal judgements and decisions	2%	17%	82%
I can identify examples of when doctors use heuristic and systematic strategies to make clinical judgements and decisions	2%	17%	82%
I have a better understanding of how others can affect my judgements and choices	17%	38%	46%
I have a better understanding of how my personal beliefs and experiences can affect my judgements and choices	19%	35%	46%
I feel more cautious about making clinical judgements and decisions	42%	44%	15%
I feel more confused about making clinical judgements and decisions	73%	20.8%	6%
I do not feel I have benefitted from the tutorial	54%	23%	23%
The tutorial taught me nothing new	57%	27%	17%
I think some formal teaching on how people make judgements and decisions should be included in medical training	12%	31%	57%

Table 27: Views about the relevance of the tutorial.

Over half of the sample stated they did not have a better understanding of how others, their personal beliefs and experiences can affect their judgements and choices. In terms of understanding the application of the tutorial content, almost the entire sample believed they could identify examples of heuristic and systematic strategies to make personal and clinical judgements and decisions. The majority of people did not feel more confused about making clinical judgements and decisions in practice after the tutorial nor did they feel more cautious, although over one third of the sample was ambivalent about feeling cautious (Table 27).

5.5.3. Clinical Judgement

Clinical judgements were measured in terms of the accuracy of diagnosis, certainty of diagnosis and how doctors reasoned about their judgement. In the following sections, the results of the McNemar tests have been reported where possible. There were occasions when the test did not compute any values. The McNemar test is calculated using the values in the diagonal of a contingency table. For the McNemar test to be possible, the data should be arranged in a symmetrical contingency table, such as a 2×2 or 3×3 table etc. (Agresti 1990). In this study, some of the cross-tabulations were not arranged symmetrically, therefore the test was not conducted. In these cases, the results of the cross-tabulations were described. Appendix 8.18 shows examples of cross-tabulations and results of statistical tests as SPSS output.

5.5.3.1. Diagnostic accuracy

Table 28 shows the proportions of correct diagnoses before and after the tutorial, in response to each clinical vignette. The majority of the sample correctly diagnosed asthma and croup before and after the tutorial.

Scenario type	Proportion of sample correct before tutorial	Proportion of sample correct after tutorial
Asthma scenario	83%	90%
Croup scenario	98%	96%
Diabetes scenario	0	2%
Angina	65%	75%

Table 28: Accuracy of diagnosis before and after tutorial.

Nobody correctly diagnosed diabetes except one person after the tutorial. Almost the entire sample misdiagnosed the patient before and after the intervention. The majority of people thought the patient had consumed alcohol excessively before and after the tutorial, 73%

and 75% respectively. The rest of the sample thought that the patient had sustained a head injury. Angina was correctly diagnosed by more people after than before the tutorial. Ten people made the same misdiagnosis after the intervention as they did before i.e. they did not change their previous incorrect responses. Of those that did change their responses after the intervention, six of them were changed to the correct diagnosis and one was changed from the correct to an incorrect diagnosis. There were no significant differences between the diagnostic judgements people made before compared to after the intervention ($\chi^2 = 4.33$, df = 2, p = 0.12).

5.5.3.2. Diagnostic certainty

Overall, the majority of people remained at the same level of certainty about their diagnoses before and after the tutorial on all scenarios (Table 29). However, the proportions of people that became more certain of their diagnoses after the tutorial was higher than those that became less certain (Table 29). The majority of people were fairly certain that their diagnosis on the asthma and croup scenario was correct, at both times.

Scenario type	Decreased certainty	Stayed same	Increased certainty
Asthma scenario	8%	73%	19%
Croup scenario	15%	67%	19%
Diabetes scenario	10%	75%	15%
Angina scenario	13%	65%	23%

Table 29: Change in certainty of diagnosis after tutorial.

There were no significant differences in people's certainty about their diagnosis on the asthma and croup scenario before and after the interventions (z = −1.615, p = 0.11; z = −0.688, p = 0.49. The sample was least certain about their diagnosis on the diabetes scenario. Over 50% of the sample was not at all certain about their chosen diagnosis before and after the intervention. There were no significant differences in people's certainty about their diagnosis on this scenario, before and after the interventions (z = −0.577, p = 0.56). People's certainty was most varied on the angina scenario, but overall most were uncertain about their diagnosis. Furthermore, before the tutorial the majority of the sample, 44%, said they were not at all certain about their diagnosis. Whereas after the tutorial a majority of 40% had slightly increased their certainty to somewhat certain. The result of the Wilcoxon test indicated that there was no significant differences in people's certainty about their diagnosis on the angina scenario, before and after the tutorial (z = −0.894, p = 0.37).

Scenario type	Correct + certain	Correct + uncertain	Incorrect + certain	Incorrect + uncertain
Asthma scenario	Before 40% After 58%	Before 42% After 33%	Before 13% After 4%	Before 2% After 4%
Croup scenario	Before 67% After 71%	Before 31% After 25%	Before 0% After 0%	Before 2% After 2%
Diabetes scenario	Before 0% After 0%	Before 0% After 2%	Before 6% After 6%	Before 94% After 92%
Angina scenario	Before 19% After 23%	Before 46% After 52%	Before 4% After 2%	Before 29% After 19%

Table 30: Comparing accuracy with certainty of diagnosis.

It was useful to know whether some people become more certain about a correct diagnosis after the tutorial, or more certain about a wrong diagnosis. Further analysis was conducted to determine how the accuracy and participants' certainty of diagnosis interacted (Table 30). Overall, the proportion of people who misdiagnosed in each scenario and did not recognise their error, was small i.e. incorrect and certain. Most people were correct about their diagnosis on the asthma and croup scenarios and were confident that they had diagnosed correctly. Further, the proportion of people who diagnosed asthma and croup correctly, and were certain of their diagnosis, increased after the tutorial. On the diabetes scenario almost the entire sample chose an incorrect diagnosis, but they were not confident that they had diagnosed correctly before or after the tutorial. Whereas on the angina scenario, most people chose the correct diagnosis but were uncertain that their judgement was correct. Further, the number of people that chose the correct diagnosis but remained uncertain, increased after the tutorial.

5.5.3.3. Excluding or diagnosing other conditions before and after intervention

For all scenarios, most people wanted to consider multiple conditions alongside their chosen diagnosis, before and after the tutorial. Further, the proportions of people that wished to consider multiple conditions increased after the tutorial, in response to all scenarios. The majority was largest on the diabetes scenario, where 83% before and 85% after the tutorial wanted to consider other conditions alongside their diagnosis. On the asthma scenario, the second commonest response was to consider panic attack 17% selected this option before the tutorial and 19% selected it after. Whereas on the croup scenario, 21% of the sample preferred not to consider any of the stated options alongside their diagnosis. The proportion of people that chose this option decreased from 21% to 10% after the tutorial.

5.5.3.4. Rating the influence of each aspect of information

On the asthma scenario, participants were asked to rate how influential the following aspects of information were on their chosen diagnosis; (i) it was the patient's first week at university, (ii) a history of asthma, (iii) signs of wheezing and shortness of breath and (iv) absence of chest pain and illness. There was a consensus that signs of wheezing and shortness of breath was the most influential aspect of information on participants' chosen diagnosis. There was no consensus about the aspect of information considered to be the least influential. Most people stated that the other three aspects of information had some influence on their diagnosis. A history of asthma was considered to have some influence by 54% of the sample, but this decreased to 35% after the tutorial. That it was the patient's first week at university had some influence on 35% of the sample before and 38% after the intervention. The absence of chest pain and illness had some influence on 33% of the sample and this increased to 50% after the tutorial. There were no significant differences in participants' responses about each aspect of information, before compared to after the intervention. The results of the Wilcoxon tests for each aspect of information were as follows; it was the patient's first week at university ($z = -0.128$, $p = 0.90$), a history of asthma ($z = -0.202$, $p = 0.84$), signs of wheezing and shortness of breath ($z = -0.144$, $p = 0.89$) and absence of chest pain and illness ($z = -0.1528$, $p = 0.13$).

On the croup scenario, participants were asked to rate how influential the following aspects of information were on their chosen diagnosis: (i) age of patient, (ii) fever, (iii) barking cough, and (iv) absence of other symptoms. There was a consensus that a barking cough influenced their diagnosis the most, however there was little agreement on the least influential aspect of information. Before the tutorial, 79% of the sample considered a barking cough as the most influential on their diagnosis but this decreased to 71% after the intervention. The remainder of the sample considered a barking cough had little influence on their diagnosis. High importance was also attributed to the age of the patient. Before the intervention, 88% of the sample stated that the age of the child had some or most influence on their diagnosis. After the tutorial, these two options were chosen by 79% of the sample indicating age was an important factor. Presence of fever was not considered to be as influential. Less than 30% stated that fever was the most influential on their diagnosis on both occasions. The absence of other symptoms elicited the most varied responses. On both occasions, approximately two thirds of the sample stated that the absence of other symptoms had some or most influence on their diagnosis. The remaining third thought it had a minimal effect on their diagnosis. The results of the Wilcoxon tests on each aspect of information were as follows; age of patient ($z = -1.615$, $p = 0.11$), fever ($z = -1.069$, $p = 0.29$), barking cough ($z = -1.414$, $p = 0.16$), absence of other symptoms ($z = -1.604$, $p = 0.11$).

On the diabetes scenario, participants were asked to rate how influential the following aspects of information were on their chosen diagnosis; (i) patient found slumped on library steps, (ii) his unkempt appearance, (iii) not fully conscious, (iv) confused state. The general trend was that participants considered all four aspects of information were influential in their chosen diagnoses. There was no aspect of information that was rated as the least influential, but there was a clear consensus that the semi-consciousness of the patient influenced choice of diagnosis the most. Before the tutorial, 77% of the sample rated the patient not being fully conscious had some or most influence on their diagnosis. This increased to 85% after the intervention. Forty-two percent of the sample changed their responses after the tutorial. That the patient was not fully conscious gained more importance for 35% of the sample after the tutorial. For this aspect of information, the Wilcoxon test indicated participants' ratings differed significantly after the tutorial compared to before ($z = -3.145$, $p = 0.02$). The confused state of the patient was also considered important and a similar pattern emerged to the above results Before the tutorial, 73% stated that the patient's confused state had some or most influence on their diagnosis, and this increased to 77% after the tutorial. Again, 42% of the sample also changed their responses after the tutorial. For the majority, a confused state influenced the diagnosis more after the tutorial. The Wilcoxon test indicated that participants' ratings for this aspect of information also differed significantly after the tutorial compared to before (-2.488, $p = 0.13$). The unkempt appearance of the patient also influenced participants' diagnoses. Around two thirds of the sample rated this to have had some influence or most influence on their chosen diagnosis. Those who rated it as the most influential aspect of information increased from 31% to 38% after the tutorial. However, there were no significant differences in these ratings after the tutorial compared to before ($z = -0.054$, $p = 0.96$). That the patient was found slumped on the library steps was less influential than other aspects of information. Before the tutorial, 63% of the sample stated it had some or most influence on their diagnosis, but only 38% chose these options after the tutorial. There were no significant differences in these ratings after the tutorial compared to before ($z = -0.386$, $p = 0.70$).

On the angina scenario, participants were asked to rate how influential the following aspects of information were on their chosen diagnosis: (i) age of patient, (ii) his active lifestyle and physical fitness,

(iii) chest pain even at rest, (iv) no history of heart problems. There was no clear consensus regarding the aspect of information considered to be the least influential on participants' diagnoses.. However, just over 50% of the sample believed that the patient's lack of history with heart problems was not influential on their diagnoses. This was the case before and after the tutorial. The Wilcoxon test revealed there were no significant differences in judgements about this aspect of information, before compared to after the tutorial ($z = -1.189$, $p = 0.235$). There was a clear consensus on the most influential aspect of information. Around 87% of the sample thought that chest pain even at rest influenced their diagnosis before and after the tutorial. Furthermore the number of people that rated it as the most influential aspect of information increased from 58% to 77% after the tutorial. There was a significant difference in ratings about chest pain at rest, before compared to after the tutorial ($z = -2.50$, $p = 0.01$). The age of the patient was also quite influential on most peoples' diagnosis. On both occasions, over 68% rated that it had some or most influence on their diagnosis. However there were no significant differences between responses before compared to after the tutorial ($z = -0.44$, $p = 0.66$). There was a similar pattern of responses towards the active lifestyle and physical fitness of the patient. This information was also influential for over 60% of the sample before and after the tutorial. On this occasion the Wilcoxon test revealed a near significant difference before compared to after the tutorial ($z = -0.554$, $p = 0.58$).

5.6. Discussion

This study reports a quasi-experiment that evaluated the feasibility of delivering an online tutorial to enhance doctors' decision literacy. All of the participants scored well on the knowledge test, indicating they did not have problems understanding the tutorial material. Despite this, most people felt unsure as to whether the tutorial had given them better knowledge about making judgements and decisions. In particular, they were unclear about the factors that can affect their judgements and choices. The sample did think the tutorial information was relevant to them as doctors and were receptive to the inclusion of this type of training into medical education. The intervention was not enough to significantly improve clinical judgements, although overall most participants' judgements were accurate and sensible on three out of four scenarios. There was evidence of biased thinking about the diabetes scenario. Everyone judged the patient to be drunk and was influenced by information that was the least relevant to the diagnosis. Most people were confident about the accuracy of their diagnosis on the asthma and croup scenarios, but not on the diabetes and angina scenario. There was evidence that the majority were considering a range of conditions alongside their chosen diagnoses. The following discussion explains the findings in more detail, and is organised in such a way to answer the study's research questions.

5.6.1. Does an intervention about basic decision science improve doctors' knowledge of how they make judgements and decisions?

An online tutorial about basic decision science was enough to impart knowledge to doctors about how they make judgements and decisions. The results of the knowledge test showed that the majority of participants understood the concepts that were introduced in the tutorial well. This was an encouraging find, given that it was unlikely the sample would have had previous knowledge about information processing. However, it is important to acknowledge that previous knowledge about the decision sciences concepts was not actually assessed. Therefore, it is possible that previous knowledge may account for participants' high scores on this knowledge test. Further, the format of the knowledge test was limited in that each question was associated with the same 6 response options that appeared in the same order (see Appendix 8.15). The disadvantage of this format is that participants may have been able to guess which response was most likely correct based on the previous answers they had chosen. Therefore, bias was potentially introduced this way and the findings should be interpreted with these limitations in mind. It is still likely that a brief online tutorial about basic decision science was enough to impart knowledge to improve doctors' decision literacy.

This finding would be in line with those interventions reviewed in Chapter 2 that reported participants had gained better knowledge about how doctors employ heuristics and how they can lead to biased judgements (Wolf et al. 1988; Hershberger et al. 1995; Round 1999). The findings from this study and previous studies indicate that it is possible to translate information from the decision sciences in a meaningful way to doctors.

As well as an objective test of knowledge through MCQs, knowledge gained was measured subjectively by asking participants their views towards the tutorial. Almost everyone felt they had been taught new information and agreed they could identify examples of decisions that had been made using a heuristic and systematic strategy. These views provided some confirmation of the high scores on the MCQ test. A similar finding was reported by Abraham (2004). Feedback from participants indicated that the intervention had increased their knowledge and understanding about critical thinking skills. However in this study, the sample did not feel they had developed better understanding of how people make judgements and decisions. Perhaps this was due to the wording of the item. If the word 'better' had been eliminated participants may have agreed more with a statement like 'I have developed an understanding of how people make decisions'. Another possible reason was that there was no feedback given to participants about their performance on any of the tasks. Feedback has been shown to be an important aspect of successful learning (King 1999; Chowdhury & Kalu 2004; Race 2005). Participants did not know their scores on the knowledge test, but if they had known then they may have responded to this item differently. They also felt unclear about the impact that others, personal beliefs and experiences had on their judgements and choices. It is likely that more detailed information and more examples were required that explained how these factors can influence their judgements and decisions than what was presented in the tutorial.

5.6.2. Can the intervention be delivered in a way that is relevant to doctors?

The majority of the sample recognised the relevance the intervention had on their work as doctors. They found the topic interesting and the intervention helped them think about their own judgements and decision making. Further, over half of the sample agreed that doctors should receive some formal training to improve their decision literacy as part of their training. There is evidence that people respond favourably to web based instruction (Chumley-Jones et al. 2002; Atreja, Mehta, Jain, Harris, Ishwaran, Avital and Fishleder 2008). In one review of web based instruction in continuing medical education, the majority of studies reported that participants had positive attitudes towards the interventions (Wutoh, Boren and Balas 2004). Further, they were as effective in imparting knowledge as traditional formats in continuing medical education (Wutoh et al. 2004). The present study also found that the online intervention was effective at imparting knowledge about decision sciences concepts to doctors. However, participants' reaction to web based training in these above studies was evaluated in terms of enthusiasm, satisfaction with learning experience and not relevance of the learning material towards the applied context. Relevance of the material was evaluated in a similar study reviewed in Chapter 2 (Servais et al. 2006). This intervention delivered a web based intervention with problem solving scenarios to improve surgical decision making. The vast majority of participants believed that the online intervention was an effective mode of teaching clinical decision making and they would use similar teaching modules in future if they were accessible over the Internet (Servais et al. 2006). These positive evaluations indicate participants did find the material and delivery mode relevant to their clinical work. That participants in the present study did find the intervention relevant to their work as doctors, was encouraging. Relevance of the learning material is an important factor in determining whether people are motivated to learn and apply their learning (Noe & Schmitt 1986).

Learning is thought to be optimised when principles are stated explicitly and plenty of relevant examples are given to clarify the principles. Evidence has found that when participants were given a concept to learn, they did not make use of it unless relevant examples were presented as well (Anderson 1987; Ross 1987; Ross & Kennedy 1990). In this study, it was anticipated that the decision making concepts were likely to be new information to the sample. The author ensured that concepts were explained using multiple examples from the non-clinical and clinical context. It is likely that the examples presented in the tutorial helped participants understand the relevance the decision sciences has on informing their clinical practice. People are usually educated about heuristics from the perspective that they can lead to poor judgements and decisions. Most studies focus on providing examples to people of heuristics that have failed (Fischhoff 1975; Wolf et al. 1988; Hershberger et al. 1995). The limitation with this approach is that people may not gain the understanding that people routinely use heuristics because they mostly lead to accurate judgements and good decisions. This study sought to give participants a more complete understanding of heuristic processing by (i) presenting examples of times when heuristics can lead to good as well as poor judgements and (ii) presenting information about systematic processing and comparing it with heuristic processing. Shepperd & Koch (2005) have recommended that any attempts to educate people about the use of heuristics should not neglect providing examples of heuristics that lead to good judgements. They found that if they gave students examples of the representative heuristic (see Section 1.6) leading to only poor judgements, it hampered their understanding of the heuristic. This group showed even less understanding than those who were not given any examples to explain the heuristic. Students that were shown examples of the heuristic leading to good and poor judgements, showed the best understanding (Shepperd & Koch 2005).

5.6.3. Does the intervention improve doctors' clinical judgements?

The intervention was not enough to improve doctors' clinical judgements about diagnosis. Overall the vast majority of people made the same judgements on the problem solving tasks before and after the tutorial. However, these were mostly accurate judgements from the start. The majority of people correctly diagnosed asthma, croup and angina before and after the tutorial. A positive find was that the majority of people correctly diagnosed angina in the forest ranger with chest pain before and after the tutorial. It was expected that fewer people would get this diagnosis correct at least before the tutorial, as it described an atypical presentation of angina in a seemingly young, fit and healthy man. It is likely that participants had been taught to always consider angina if a patient complains of chest pain even at rest. They may have used a method of ruling out the worst case scenario (Croskerry 2002) which in this case was angina. While the intervention was not associated with an overall positive effect, it was encouraging that it was not associated with a negative effect on doctors' judgements. Participants did not change their correct diagnoses to incorrect diagnoses after the tutorial. This indicates that learning basic information about the decision sciences did not have a detrimental impact on clinical judgements. Further, hardly anyone reported feeling more confused about making clinical judgements and decisions in practice. None of the studies reviewed in Chapter 2 evaluated whether their intervention had a negative impact on clinical judgements and/or decisions.

There was evidence of biased reasoning on the diabetes scenario. All except one doctor inaccurately judged the diabetes patient to be drunk, before and after the tutorial. This was not just an incorrect diagnosis, but it suggested that participants' reasoning about the situation had been influenced by the misleading aspects of the scenario. For example, information such as an unkempt appearance and the patient found slumped on the steps of a public library biased their judgement about the cause of the patient's condition. That everybody thought the patient had consumed too much alcohol, suggests doctors were influenced by a prototypical typical appearance of a drunken man. This is evidence that participants had over relied on the confirmation heuristic when thinking about this clinical scenario. That is, their judgements were based on how representative the clinical information was about a drunken man (Tversky & Kahneman 1974; Croskerry 2000). The information presented in the tutorial did not help doctors think differently about this particular scenario, for example that the cause of the patient's symptoms could be an atypical presentation of an illness.

As the intervention did not improve clinical judgements, these findings do not support the view that doctors can avoid the influence of bias through education about heuristics (Arnoult & Anderson 1988; Croskerry 2000; Croskerry 2003). The systematic review in Chapter 2 found a few interventions that educated doctors about heuristics and bias were associated with better diagnostic judgements (Wolf et al. 1988; Gruppen et al. 1994; Hershberger et al. 1995; Round 1999; Abraham et al. 2004). There may be a few reasons for why the same was not found of the present study. Overall, most people made good clinical judgements in response to all but one of the problem solving tasks, before the tutorial. Changes in responses that were better than previous answers were so few that they did not result in significant differences. There was only room for significant improvement in response to how they solved the diabetes scenario. Further, the study's sample size was fairly small (N = 48). A small sample size is one factor that compromises the power of a statistical test i.e. the probability of rejecting the null hypothesis when it is indeed false (Cohen 1988). This means that the statistical power in this study may not have been sufficient to detect a true effect of the intervention.

It may be that this type of intervention genuinely has no effect on doctors' clinical judgements. That is, teaching doctors about basic decision science is not an effective way of enhancing clinical judgements. However, this is unlikely as the intervention was based on good rationale; the systematic review found that decision literacy interventions were effective i.e. those that taught students and doctors about critical thinking or the reasoning processes they use to solve clinical problems (Cebul et al. 1984; Wolf et al. 1988; Hershberger et al. 1995; Round 1999; Beullens et al. 2006). Further, the study took on board the suggestions of other authors that doctors should be educated about their use of heuristics from an approach that emphasises they are routinely used because they are helpful (Croskerry 2002; Croskerry 2003; Eva & Norman 2005; Shepperd & Koch 2005). It is more likely that the intervention was delivered in a format that was not suitable to impact on doctors' clinical judgements. The online tutorial was perhaps too brief a delivery mode for educating doctors about decision making concepts and information processing. It may have been more suitable to have delivered this information in a series of lectures or tutorials by an expert in the decision sciences. Doctors may have gained more benefit from having contact with a teacher and their peers in order to discuss the information about decision making and receive feedback on their performance on problem solving tasks. This mode of delivery is thought to facilitate deep learning (Spencer & Jordan 1999) and there is reason to believe that this format would have been more likely to improve clinical judgements. All of the interventions that delivered some education in critical thinking or information processing and were associated with an improvement in clinical judgements, delivered their teaching in a lecture or tutorial format (Cebul et al. 1984; Wolf et al. 1988; Round 1999; Abraham et al. 2004; Beullens et al. 2006).

5.6.4. Implications of findings to medical education and practice

Medical educators should consider introducing some formal teaching about the decision sciences for trainees. It is likely that doctors will be receptive to it if they understand the relevance it has on their clinical practice. It is possible to translate information from the decision sciences in a meaningful way to doctors. In this study people gained knowledge about information processing. This means that doctors can be taught to understand how the generic principles of judgement and decision making are relevant to their clinical practice. It is unlikely that online education about the decision sciences will be enough on its own to improve clinical judgements and/or decisions in practice. Interactive forms of teaching are necessary to enhance deeper learning so participants can recognise when they need to think differently about a clinical problem. Web based training is a satisfactory mode of learning when it supplements traditional forms of teaching. For instance, doctors reportedly find it useful and enjoyable to practise solving clinical problems using online scenarios, that are often more graphically rich than paper based scenarios (Servais et al. 2006; Hedrick & Young 2008).

Doctors displayed biased thinking and this was not corrected after they had been given information about heuristics and the errors they can lead to i.e. they did not transfer their learning from the tutorial to solve this problem. It is reasonable to assume that participants would have made similar judgements in practice as all of the clinical vignettes were based on real clinical situations. This highlights a training need for doctors that they become more reflective about their reasoning processes and aware of the factors that influence their judgements and decisions. In particular there was evidence that participants over relied on the confirmation heuristic. As this is one of the most commonly used rules of thumb by all individuals, medical educators should be cautious to place equal emphasis on teaching trainees to recognise patterns of atypical as well as typical signs and symptoms of medical conditions (Croskerry 2002).

Little is known about how clinical reasoning interventions impact on clinical practice. Interventions that seek to raise doctors' awareness about judgement and decision making processes have been criticised on the basis that learning will not be transferred to the clinical context to improve practice (Gick & Hollyoak 1983; Norman et al. 1995; Nendaz & Bordage 2002; Schuwirth 2002; Norman 2005; Eva & Norman 2005; Norman & Eva 2010). At this stage, it is more important to establish that clinical reasoning interventions do not have a detrimental effect on real world clinical judgements and decisions. This may happen if doctors became confused about how they make decisions. In this study, the intervention did not harm the correct judgements doctors made. Most interventions reviewed in Chapter 2 did not evaluate this aspect and further research is needed in this area.

5.6.5. Critique of study

The strengths of this study include a successful translation of basic information from the decision sciences in a way that was relevant and interesting to doctors. It avoided limitations of other studies that risk creating the false impression of a fundamental problem with the way doctors use heuristics to make judgements and decisions, and the need to avoid using heuristics. This study presented the use of heuristics in its broader context as one mode of processing information. It emphasised the importance of these mechanisms that simplify information processing and the effective results they produce most of the time. A systematic method was presented as an alternative way that doctors can use to solve clinical problems when heuristics are inappropriate. It was encouraging that participants reported that the tutorial information helped them to reflect on their own decision making and that this knowledge did not confuse them about how they should make judgements and decisions in practice.

From a review of the evidence (Chapter 2) it was unclear as to which component parts of interventions are associated with the enhancement of doctors' decision making skills. However, a summary of good practice for designing educational interventions was provided (page 72). The strengths of this study include explicit use of theory to inform the design of the intervention. Further, as this theory came from the decision sciences participants were given information about generic aspects of decision making that affect all individuals as well as specific information about their clinical judgements and decisions. Most interventions reviewed in Chapter 2 were not informed by theory and those that were focused on specific aspects of doctors' clinical reasoning. The types of measures were designed specifically for the present study in order to ensure they suitable matched the aims and objectives. This was noted as a limitation of a few of the interventions reviewed in Chapter 2, where an established measure had been used that was not well suited to the study aims and objectives. Further, the advice of Miller (1999) was adhered to by not only including a measure of knowledge, but the application of knowledge to establish how doctor's actual clinical judgements may be in practice.

Some limitations should be considered. The sample size was small and only included qualified doctors who were postgraduate students at the University of Leeds. The findings do not necessarily reflect the views and clinical judgements of medical students or doctors who are not engaged in postgraduate studies. The intervention needs to be evaluated with a variety of different samples and in large numbers to gain a better understanding of whether an online intervention about basic decision science is feasible. However, the sample did at least consist of doctors from different levels of experience who worked in different clinical specialties. Delivering the intervention online meant that participants did not have an opportunity to discuss or ask questions about the problem solving tasks or tutorial material with the author or their peers. Further, the software used to design the study did not allow feedback to be given to participants on their performance after they had completed the study. This was not ideal as the provision of constructive feedback is an important aspect of successful learning (King 1999; Race 2005). Perhaps if they had been given feedback on aspects they answered well and areas for improvement then the intervention may have had more impact on their clinical judgements, particularly with regard to the diabetes scenario. It is not known whether participants will retain the knowledge they gained from the tutorial and/or go onto apply it in practice. This could be addressed in future work in order to understand if the intervention is associated with any long term benefit.

5.7. Summary

To summarise, this study evaluated the feasibility of delivering brief online teaching about basic decision sciences concepts to doctors, in order to improve their decision literacy. It is possible to impart knowledge about the decision sciences in a way that is meaningful and relevant to doctors. They were receptive to learning about the science behind their judgement and decision making. The online intervention was not enough to impact on their clinical judgements. It is anticipated that clinical judgements can be improved if the decision science information is delivered by an expert over a longer duration in an interactive teaching format.

6.

Discussion

The purpose of this thesis was to (i) develop a theory-based intervention to facilitate doctors' clinical judgements and decisions and (ii) evaluate the feasibility to integrate it within medical education. The objectives were to:

- identify the evidence of interventions designed to enhance doctors' clinical reasoning;
- evaluate the effectiveness of those interventions;
- describe how interventions have been incorporated into medical curricula;
- explore doctors' views and experiences about making clinical judgements and decisions effectively;
- make recommendations for the design and implementation of future clinical reasoning training for doctors.

The aims and objectives were fulfilled by carrying out four empirical studies reported in Chapters 2, 3, 4 and 5. Evidence of training courses to enhance doctors' clinical judgements and decisions was synthesised in the systematic review and questionnaire studies (Chapters 2 and 3). The main findings were that few formal courses have been designed to train doctors in clinical reasoning. Of the courses identified, most of them had not been included into medical curricula. This indicates improving doctors' clinical reasoning through formal training has not gained importance in medical education. There was little agreement between medical educators about how to improve doctors' clinical judgements. However, training to improve the outcome of decisions remained a more common approach than training to improve the decision process.

Doctors were interviewed about their perceptions of making clinical judgements and decisions effectively (Chapter 4). The sample had limited awareness about their own reasoning processes and did not feel there was an ideal way to make effective clinical judgements or decisions. There was some awareness that a diverse array of cues influenced their thought processes but little awareness about the occurrence of biased thinking. It was found that doctors in the sample believed that clinical reasoning improved primarily through years of repeated experience in practice.

The findings from the systematic review, questionnaire and interview studies informed the development of an online decision literacy intervention (Chapter 5). A short tutorial was delivered that introduced the basic science behind doctors' judgement and decision making. The intervention was enough to impart knowledge to doctors about their reasoning processes, and was received well by participants. It did not impact on their clinical judgements, but this was likely to be due to methodological limitations rather than an ineffective intervention. It is anticipated that the inclusion of the decision sciences in medical curricula will be feasible and beneficial to doctors.

This chapter discusses the contributions of this thesis to the clinical reasoning literature, medical education and training as well as directions for future research. It ends with a critique of the thesis.

6.1. Implications of findings to the clinical reasoning literature

A distinction between judgements and decisions is clear and more accurate from a decision sciences perspective. In this thesis, an attempt was made at distinguishing clinical judgements from clinical decisions. However, it was difficult to apply this same distinction to the clinical reasoning context because it is not clear at which point clinical judgements become decisions and vice versa. This thesis imposed a separation between forming an opinion about a diagnosis (judgement) and making a choice about how to manage a patient (decision). There are a few advantages of separating clinical judgements from clinical decisions in this way. First, the aims and objectives of studies become clearer. The current use of interchangeable terminology makes it difficult to understand what research problem medical educators are trying to address. For example, whether the aim is to improve the accuracy of a diagnosis or problem solving ability. This unclear use of terms was reflected in Chapters 2 and 3 which found there was little consensus between educators on how to effectively enhance doctors' clinical reasoning. To rectify this, authors should set out a clear rationale for what they mean by clinical judgement, clinical decision, problem solving or clinical reasoning etc.

Second, a distinction between clinical judgements and clinical decisions was useful in highlighting where gaps exist in our knowledge about improving doctors' clinical reasoning. It was clear that the vast majority of research focused on one aspect of clinical reasoning, understanding the processes doctors employ to make a diagnosis and factors that affect diagnosis (Elstein 1972; Barrows 1982; Groen & Patel 1985; Croskerry 2002; Norman 2005). Consequently, most training courses in clinical reasoning were aimed at improving doctors' diagnostic accuracy via the application of Bayes theorem and decision analysis to clinical problems (Elstein 1981; Elstein 1985). There has been little change in the approach that medical educators take to improve doctors' clinical reasoning. That is, the application of statistical decision theories is still preferred as a means of educating doctors. Little is known about how doctors choose particular courses of action to manage a patient's condition and the factors that influence those choices (Norman 2005). For this to be rectified, medical educators should be made aware that research on doctors' clinical reasoning is almost entirely focused on understanding clinical diagnosis and that this is only one aspect of clinical reasoning.

The emphasis on improving diagnosis may be due to the feasibility of conducting research. The accuracy of a diagnosis is an easily measured outcome in experimental studies; it is either correct or incorrect. Alternatively, it may reflect a commonly held view that a good decision is one that leads to a good outcome, such as the correct diagnosis (Frisch 1994; Sox 1999; Baron 2008; Bekker 2010). An important find from this thesis was that decision literacy interventions, although few in number, were effective. These involved explicit use of decision theory to enhance doctors' awareness of their reasoning processes and critical thinking skills (Wolf 1988; Hershberger 1995; Round 1999; Abraham 2004). These findings suggest that improving doctors reasoning about clinical problems offers a promising alternative way to enhance the quality of clinical judgements and decisions. The findings also shed light on ways that doctors can be helped to make well-reasoned choices. The view that a good decision is one that is reasoned well, has received little attention in the clinical reasoning literature (Croskerry 2000; Croskerry 2002). This thesis suggests that training doctors to understand their own reasoning processes can potentially improve their judgements and decisions in clinical practice (Chapter 2). When interviewed, doctors reported that they preferred the quality of their judgements and decisions are evaluated by the steps they had taken to reach them (Chapter 4). Furthermore, doctors were receptive to learning about the basic science behind their decision processes (Chapter 5). Overall, these findings substantiate the view that training doctors to reason well is a worthwhile goal.

6.2. Implications of findings to medical education and training

The findings from this thesis can be used to improve the design and delivery of future courses that seek to educate doctors about their clinical reasoning. The following sections make suggestions about how future training in clinical reasoning could be delivered in terms of content and format, who it should be aimed at and where to include it in the medical curricula.

Medical educators have continued to hold the opinion that doctors should learn how to apply statistical decision theories to solve clinical problems (Croskerry 2005). The majority of courses reviewed sought to teach doctors how to apply Bayes Theorem to their clinical diagnosis and/or decision analysis to their treatment choices. In practice, it is reasonable to expect that doctors will not always implement decision making techniques such as Bayes theorem, decision trees or clinical guidelines. For instance, evidence demonstrates it is difficult to get doctors to implement best practice guidelines (Greco 1993; Cabana 1999; Rello 2002). Teaching doctors to become decision literate could prove to be a more practical way of enhancing the quality of clinical judgements and decisions. This involves helping doctors develop an insight and ability to critically evaluate their own reasoning processes.

One of the reasons that a decision literacy approach to training may be effective is because it integrates the generic aspects of human judgement and decision making with the applied clinical context. Referring back to the good practice guidance (see page 72) a good course in clinical reasoning would have an explicit theoretical basis. The decision sciences are a useful theoretical basis that provides knowledge of generic decision making principles that are applicable to all people. When this information is combined with health and illness examples of judgements and decisions, the evidence and theories from the decisions sciences becomes relevant to the clinical context. This combination means that doctors can learn about the similarities between how all individuals generally and doctors specifically make judgements and decisions, and the factors that influence their reasoning. Findings from the quasi-experiment (Chapter 5) in this thesis demonstrated that the information processing framework was a useful way of teaching doctors to understand the basic science that underlies their personal and clinical decisions. Doctors gained a basic understanding of how the generic principles of decision making apply to their clinical practice. This knowledge could help doctors gain an ability to monitor their own thinking, and thereby become more reflective practitioners (Kassirer 1991; Croskerry 2000; Graber 2003; Graber 2005; GMC 2009; Graber 2009).

The content of such courses may include understanding how people use a variety of cognitive shortcuts (heuristics) and the effects they can have on clinical judgements and decisions (Wolf 1988; Hershberger 1995; Round 1999; Beullens 2006). A comprehensive list of the different types of heuristics that doctors use has been developed that could be used as a teaching resource (Croskerry 2002; Croskerry 2003). It is important to provide examples of when heuristics are used effectively and when they may lead to an incorrect and/or poor decision. This may help doctors develop an insight into how errors occur in their reasoning and potentially how they may be avoided in future.

Further to the good practice guidance summarised on page 72, it is important that the outcome measures are clarified and that assessment methods are appropriate. A more accurate use of terms will clarify the outcomes that are being measured, e.g. whether a judgement or decision is being measured. Future interventions should also employ assessment methods targeting not only knowledge components such as recall tests but methods that demonstrate what trainees will do when faced with real clinical problems (Miller 1990). There should also be some attempt at assessing the long-term impact of training and transfer of learning to clinical practice. In terms of how to deliver teaching about the decision sciences, an interactive series of tutorials are expected to be more suitable than online teaching. An interactive set of tutorials would allow for timely, constructive feedback about participants learning. The limitations of the online tutorial reported in Chapter 5 were that information could only be conveyed briefly and there was no face to face interaction between the educator and learners. This mode of delivery may have inhibited the effectiveness of the intervention.

A final aspect of good practice in clinical reasoning training is to plan carefully where it could be integrated into medical curriculum. From this thesis, educators raised two main issues for consideration:. (i) whether it should be taught as part of a module on a broader topic or as a module itself and (ii) whether it should be introduced before or after clinical experience is gained. Many of the educators surveyed believed there was a close link between communication skills and improving clinical reasoning. Consequently, they reported that clinical reasoning training should be integrated within modules that primarily teach doctors better communication skills. Given that an ability to reach good clinical judgements and decisions is a crucial aspect of a doctor's clinical practice, training in clinical reasoning deserves a more prominent position in medical education. This could be in the form of a series of modules devoted entirely to clinical reasoning that are integrated throughout the curriculum.

The majority of educators believed that the ideal time to deliver this training was at postgraduate level when doctors have gained sufficient clinical experience. Otherwise, an alternative option was to introduce it earlier during the later years of undergraduate training to coincide with the clinical attachment phases. A different view exists within the clinical reasoning literature that states it is not necessary to wait for medical students to develop sufficient experience of making decisions in practice and that basic principles can be introduced in the first year at undergraduate level (Croskerry 2000; Croskerry 2005; Croskerry 2009; Kassirer 2010). A few studies have supported this view by demonstrating that medical students without any clinical experience can benefit from formal training in clinical reasoning (Margolis 1982; Vollebregt 2005). It is likely that a basic introduction to the decision sciences could be introduced effectively early on in undergraduate medical training. This material would then need to be revisited later on when students have gained some clinical experience and can apply real clinical scenarios to the theory. Ideally, this material would be spiraled throughout undergraduate curricula in order to reinforce learning (Harden 2000).

6.3. Further research and future directions

The thesis has highlighted several areas of further research necessary in this subject area. Further research is needed to understand why training in clinical reasoning has made little impact in medical education. Some of the possible reasons were highlighted in this thesis, and include a firm belief that clinical reasoning skills are acquired solely at the workplace, not the classroom. Others are concerned that, despite its potential value, the curriculum is already saturated with subjects of a higher priority. Future research may want to understand the perceived barriers to clinical reasoning training better so they can be addressed if perceptions are to change.

There is considerable variation in the way decision making related terms and concepts are used amongst researchers. This compromises the quality of the clinical reasoning literature and people's understanding of how to effectively train doctors to make better judgements and decisions. An agreeable and workable language needs to be established to achieve a common understanding of what is meant by a good judgement or good decision. Studies describing future decision making courses should be explicit about whether the aim is to improve reasoning processes or the medical outcome. Medical educators should understand this important difference in approach and the implications it has on clinical practice. Achieving consistency in the use of these terms will lead to clearer aims, clearer operationalisation of variables and meaningful communication among educators about decision making training.

Understanding how doctors make treatment choices and other decisions about managing patients, has been neglected. This should be addressed in order to complement what is known about diagnosis. It is reasonable to expect that doctors employ expert methods that are different to those used to reach a diagnosis. For instance, pattern recognition is one type of method used to make a diagnosis but has little relevance to making a treatment choice. The interview study reported in Chapter 4 suggested that diagnosing a patient is reached by different processes and influenced by different factors compared to making treatment choices. To understand how to improve doctors' treatment and other patient management choices, researchers should seek to clarify the processes doctors use and the factors that influence treatment choices. Researchers should also look to the shared decision making literature and integrate any useful information into the clinical reasoning literature that informs the training of doctors in this area.

Finally, the evaluation of further training courses needs to eventually target the impact of training at the workplace. Further research needs to resolve the following main issues: (i) whether doctors apply their learning from a training course to the clinical context and (ii) what impact is there on clinical judgements and decisions in practice. The majority of educators focused on whether their courses had a positive effect on judgements and decisions. The possibility that training may worsen doctors' judgements and decisions has been overlooked. Further, it is not known whether any of the courses reviewed had a long term impact on clinical judgements and decisions. These are important aspects that must be accounted for before efforts can be made to implement any training in medical curricula. Studies that use a longitudinal design would be appropriate to address these issues.

6.4. Critique of thesis

The strengths of the thesis include a thorough exploration about how to effectively train doctors in clinical reasoning to make better judgements and decisions. It resulted in the development of a feasible online educational intervention that was based on theory and evidence. The evidence was collated over a series of empirical studies using a range of rigorous research methods. It has made a significant contribution to furthering understanding about effective and ineffective ways of training doctors in clinical reasoning and the processes that doctors use. This thesis has also demonstrated that doctors can be trained to understand how the generic principles of judgement and decision making can inform their clinical practice.

The thesis has some limitations. There is no guarantee that the evidence base reviewed in Chapters 2 and 3 represents every course in clinical reasoning. It is possible that the systematic review missed a relevant study that was not published in the English language. Further, a few medical educators may have chosen not to participate in the questionnaire study and therefore the details of these course remain unknown. Sample sizes were relatively small in the questionnaire, interview and quasi-experimental studies and a non-probability sampling method was used to recruit participants. Therefore it is not possible to claim these findings are applicable to other populations of doctors (Robson 2002). The online delivery mode of the decision literacy intervention limited the amount of information that could reasonably be conveyed. Participants may have benefitted more from the intervention if it were delivered in a format that facilitated interaction between others.

6.5. Summary

There is increasing evidence to support the value of introducing formal clinical reasoning training into medical education. This evidence suggests that there are aspects of good clinical judgement and decision making competency that are not learned through repeated clinical practice alone. In particular this thesis highlights that training to improve doctors understanding about how people think has the potential to impact on clinical reasoning. Doctors in training should be made aware that making well reasoned judgements and choices is one of their most important skills. Medical educators should develop effective training methods to improve doctors' critical thinking and reasoning processes.

7.

References

ABRAHAM, R., UPADHYA, S., TORKE, S. and RAMNARAYAN, K. (2004) Clinically oriented physiology teaching: strategy for developing critical-thinking skills in undergraduate medical students. *Advances Physiology Education*, 28, 102-104.

AGRESTI, A. (1990) *Categorical data analysis.*, Second ed., New York: Wiley.

AKICI, A., GOREN, M., AYPAK, C., TERZIOGLU, B. and OKTAY, S. (2005) Prescription audit adjunct to rational pharmacotherapy education improves prescribing skills of medical students. *European Journal of Clinical Pharmacology*, 61, 643-650.

AKICI, A., KALACA, S., GOREN, M., AKKAN, A., KARAALP, A., DEMIR, D., UGURLU, U. and OKTAY, S. (2004) Comparison of rational pharmacotherapy decision making compete of general practitioners with intern doctors. *European Journal of Clinical Pharmacology*, 64, 75-82

AKICI, A., KALAÇA, S., UGURLU, M., KARAALP, A. and ÇALI S, O. S. (2003) Impact of a short postgraduate course in rational pharmacotherapy for general practitioners. *British Journal of Clinical Pharmacology*, 57, 310-321.

ANDERSON, J. (1987) Skill acquisition: Compilation of weak method problem solutions. *Psychological Review*, 94, 192-210.

ANDERSON, J., MCEWAN, K. and HRUDEY, W. (1996) Effectiveness of notification and group edication in modifying prescribing of regulated analgesics. *Canadian Medical Association Journal*, 154, 31-39.

ANDERSON, W. & HARRIS, I. (2003) Arthur S. Elstein, Ph.D.: Skeptic, Scholar, Teacher and Mentor. *Advances in Health Sciences Education*, 8, 173–182.

ANTHONY, T. (1986) A discrepancy in objective and subjective measures of knowledge: do some medical students with learning problems delude themselves? . *Medical Education*, 20, 17-22.

ARK, T., BROOKS, L. and EVA, K. (2006) Giving learners the best of both worlds: do clinical teachers need to guard against teaching pattern recognition to novices? *Academic Medicine*, 81, 405-409.

——— (2007) The benefits of flexibility: the pedagogical value of instructions to adopt multifaceted diagnostic reasoning strategies. *Medical Education*, 41, 281-287.

ARNOULT, L. & ANDERSON , C. (1988) Identifying and reducing causal reasoning biases in clinical practice. IN TURK, D. & SALOVEY, P. (Eds.) *Reasoning, inference and judgement in clinical psychology*. Macmillan, New York.

ATREJA, A., MEHTA, N., JAIN, A., HARRIS, C., ISHWARAN, H., AVITAL, M. and FISHLEDER, A. (2008) Satisfaction with web-based training in an integrated healthcare delivery network: do age, education, computer skills and attitudes matter? *BMC Medical Education*, 8.

BARON, J. (2008) *Thinking and Deciding*, 4th ed., Cambridge University Press.

BARROW, M., MCKIMM, J. and SAMARASEKERA, D. (2010) Strategies for planning and designing medical curricula and clinical teaching. *South East Asian Journal of Medical Education*, 4, 2-8.

BARROWS, H. & FELTOVICH, P. (1987) The Clinical Reasoning Process. *Medical Education*, 21, 86-91.

BARROWS, H., NORMAN, G., NEUFELD, V. and FEIGHTNER, J. (1982) The clinical reasoning process of randomly selected physicians in general medical practice. *Clinical and Investigative Medicine*, 5, 49-56.

BARROWS, H. & TAMBLYN, R. (1980) *Problem based learning: An approach to medical education*, Springer, New York.

BECK, A. & BERGMAN, D. (1986) Using structured medical information to improve students' problem-solving performance. *Journal of Medical Education*, 61, 749-756.

BEKKER, H. (2010) The loss of reason in patient decision aid research: Do checklists damage the quality of informed choice interventions? *Patient Education and Counseling* 78, 357-364.

BELLON, J. A. & FERNANDEZ-ASENSIO, M. E. (2002) Emotional profile of physicians who interview frequent attenders. *Patient Education and Counseling*, 48, 33-41.

BENDTSEN, P., HENSING, G., EBELING, C. and SCHEDIN, A. (1999) What are the qualities of dilemmas experienced when prescribing opioids in general practice? *Pain*, 82, 89-96.

BENNER, P. (1982) From Novice to Expert. *The American Journal of Nursing*, 82, 402-407.

BEULLENS, J., STRUYF, E. and VAN DAMME, B. (2006) Diagnostic ability in relation to clinical seminars and extended-matching questions examinations. *Medical Education*, 40, 1173-1179.

BORNSTEIN, B. H., CHRISTINE EMLER, A. and CHAPMAN, G. B. (1999) Rationality in medical treatment decisions: is there a sunk-cost effect? *Social Science & Medicine*, 49, 215.

BORNSTEIN, B. H. & EMLER, A. C. (2001) Rationality in medical decision making: a review of the literature on doctors' decision-making biases. *Journal of Evaluation in Clinical Practice*, 7, 97-107.

BRAUN, V. & CLARKE, V. (2006) Using thematic analysis in psychology. *Qualitative research in psychology*, 3, 77-101.

BRENNAN TA, LEAPE LL, LAIRD NM, HEBERT L, LOCALIO AR, LAWTHERS AG, NEWHOUSE JP, WEILER PC and HIATT, HH. (1991) Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical: Practice Study 1. *New England Journal of Medicine*, 324, 370-376.

BROOKS, L., NORMAN, G. and ALLEN, S. (1991) Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General*, 120, 278-287.

BROTRNAN, A., STERN, T. and HERZOG, D. (1984) Emotional Reactions of House Officers to Patients with Anorexia Nervosa, Diabetes, and 0besity. *International Journal of Eating Disorders*, 3, 71-77.

BRYMAN, A. & CRAMER, D. (2008) *Quantitative Data Analysis with SPSS 14, 15 and 16: A Guide for Social Scientists*, Routledge.

BYRNE, A., SELLEN, A., JONES, J., AITKENHEAD, A., HUSSAIN, S., GILDER, F., SMITH, H. and RIBES, P. (2002) Effect of videotape feedback on anaesthetists' performance while managing simulated anaesthetic crises: a multicentre study. *Anaesthesia*, 57, 169-182.

CABANA, M., RAND, C., POWE, N., WU, A., WILSON, M., ABBOUD, P. and RUBIN, H. (1999) Why don't physicians follow clinical practice guidelines? A framework for improvement. *Journal of the American Medical Association*, 282, 1458-1465.

CAMPBELL, N., MURRAY, E., DARBYSHIRE, J., EMERY, J., FARMER, A., GRIFFITHS, F., GUTHRIE, B., LESTER, H., WILSON, P. and KINMONTH, A. (2007) Designing and evaluating complex interventions to improve health care. *British Medical Journal* 334, 455-459.

CARTER, B., BUTLER, D., ROGERS, J. and HOLLOWAY, R. (1993) Evaluation of physician decision making with a use of prior probabilities and a decision-analysis model. *Archive of Family Medicine*, 2, 529-534.

CEBUL, R., BECK, L. and CARROLL, J., ET AL (1984) A course in clinical decision making adaptable to diverse audiences. *Medical Decision Making*, 4, 285-296.

CHAIKEN, S. (1980) Heuristic versus systematic information processing and the use of Source Versus Message Cues in Persuasion. *Journal of Personality and Social Psychology*, 39, 752-766.

CHALMERS, I. & ALTMAN, D. (1995) *Systematic Reviews*, London, BMJ Publishing.

CHAPMAN, G. & ELSTEIN, A. (2000) Cognitive processes and biases in medical decision making. IN CHAPMAN, G. & SONNENBERG, F. (Eds.) *Decision Making in Health Care: Theory, Psychology, and Applications*. New York: Cambridge University Press.

CHARLES, C., GAFNI, A. and WHELAN, T. (1997) Shared decision-making in the medical encounter: What does it mean? (or it takes at least two to tango). *Social Science & Medicine*, 44, 681-692.

CHEN, H. T. (1990) *Theory-driven evaluations*, Sage Publications.

CHEN, S. & CHAIKEN, S. (1999) The heuristic-systematic model in its broader context. IN CHAIKEN, S. E. & TROPE, Y. E. (Eds.) *Dual-process theories in social psychology.* New York, NY, US: Guilford Press; US.

CHESSARE, J. & LIEU, T. (1998) Teaching Clinical Decision-making to Pediatric Residents in an Era of Managed Care. *Pediatrics*, 101, 762-767.

CHOPRA, V., GESINK, B., DE JONG, J., BOVILL, J., SPIERDIJK, J. and BRAND, R. (1994) Does training on an anaesthesia simulator lead to improvement in performance? . *British Journal of Anaesthesia*, 73, 293-297.

CHOWDHURY, R. R. & KALU, G. (2004) Learning to give feedback in medical education. *The Obstetrician and Gynaecologist*, 6, 243-247.

CHUMLEY-JONES, H., DOBBIE, A. and ALFORD, C. (2002) Web-based Learning: Sound Educational Method or Hype? A Review of the Evaluation Literature. *Academic Medicine*, 77.

CLARKE, J., SPEJEWSK, I. B., GERTNER, A., WEBBER, B., HAYWARD, C., SANTORA, T., WAGNER, D., BAKER, C., CHAMPION, H., FABIAN, T., LEWIS, F. J., MOORE, E., WEIGELT, J., EASTMAN, A. and BLANK-REID, C. (2000) An objective analysis of process errors in trauma resuscitations. *Academic Emergency Medicine*, 7, 1303-1310.

CODERRE, S., MANDIN, H., HARASYM, P. and FICK, G. (2003) Diagnostic reasoning strategies and diagnostic success. *Medical Education*, 37, 695–703.

COHEN, J. (1988) *Statistical poer analysis for the behavioural sciences*, Second ed., New York: Academic Press.

COOK, D. (2007) Web-based learning: pros, cons and controversies. *Clinical Medicine*, 7, 37-42.

COOK, T. & CAMPBELL, D. (1979) *Quasi-experimentation: design and analysis issues for field settings*, Chicago: Rand McNally College Pub. Co.

COOLICAN, H. (1999) *Research methods and statistics in psychology*, 3rd ed., Hodder & Stoughton.

CRD (2001) Undertaking systematic reviews of research on effectiveness. CRD's guidance for those carrying out or commissioning reviews.

CRESWELL, J. (1998) *Qualitative inquiry and research design: Choosing among five traditions.*, Thousand Oaks, CA: Sage.

CROSKERRY, P. (2000) The Cognitive Imperative Thinking about How We Think. *Academic Emergency Medicine*, 7, 1223-1231.

——— (2002) Achieving Quality in Clinical Decision Making: Cognitive Strategies and Detection of Bias. *Academic Emergency Medicine*, 9, 1184-1204.

——— (2003) The importance of cognitive errors in diagnosis and strategies to minimize them. *Academic Medicine*, 78.

——— (2005) The theory and practice of clinical decision-making. *Canadian Journal of Anesthesia*, 56, R1-R8.

——— (2009a) Critical Thinking and Reasoning in Emergency Medicine, Ch 31. IN CROSKERRY, P., CROSBY, K., SCHNEKEL, S. & WEARS, R. (Eds.) *Patient Safety in Emergency Medicine.* Lippincott Williams and Wilkins

——— (2009b) A universal model of diagnostic reasoning. *Academic Medicine*, 84, 1022–1028.

CROSKERRY, P., WEARS, R. and BINDER, L. (2000) Setting the Educational Agenda and Curriculum for Error Prevention in Emergency Medicine. *Academic Emergency Medicine*, 7, 1194-1200.

CUFF, P. A. & VANSELOW, N. A. (Eds.) (2004) *Improving medical education: enhancing the behavioural and social sciences content of medical school curricula*, Washington D.C, The National Academies Press.

DAROSA, D., ROGERS, D., WILLIAMS, R., HAUGE, L., SHERMAN, H., MURAYAMA, K., NAGLE, A. and DUNNINGTON, G. (2008) Impact of a structured skills laboratory curriculum on surgery residents' intraoperative decision-making and technical skills. *Academic Medicine*, 83, S68-71.

DAVIDOFF, F., GOODSPEED, R. and CLIVE, J. (1989) Changing test ordering behaviour: A randomized controlled trial comparing probabilistic reasoning with cost-containment education (placebo). *Medical Care*, 27, 45-58.

DE DOMBAL, F., DALLOS, V. and MCADAM, W. (1991) Can computer aided teaching packages improve clinical care in patients with acute abdominal pain? *British Medical Journal*, 302, 1495-1497.

DE VRIES, T., DANIELS, J., MULDER, C., GROOT, O., WEWERINKE, L., BARNES, K., BAKATHIR, H., HASSAN, N., VAN BORTEL, L., KRISKA, M., SANTOSO, B., SANZ, E., THOMAS, M., ZIGANSHINA, L., BEZEMER, P., VAN KAN, C., RICHIR, M. and HOGERZEIL, H. (2008) Should medical students learn to develop a personal formulary? *European Journal Clinical Pharmacology*, 64, 641-646.

DENIG, P., WITTEMAN, C. and SCHOUTEN, H. (2002) Scope and nature of prescribing decisions made by general practitioners. *Quality and safety in healthcare*, 11, 137-143.

DEVRIES, T. (1993) Presenting clinical pharmacology and therapeutics: evaluation of a problem based approach for choosing drug treatments. *British Journal of Clinical Pharmacology*, 35, 591-597.

DEVRIES, T., HENNING, R., HOGERZEIL, H., BAPNA, J., BERO, L., KAFLE, K. K., MABADEJE, A. F. B., SANTOSA, B. and SMITH, A. (1995) Impact of a short course in pharmacotherapy for undergraduate medical students: an international randomised controlled study *The Lancet*, 346, 1454-1457.

DONNELLY, M., SISSON, J. and WOOLLISCROFT, J. (1990) The reliability of a hypothesis generation and testing task. *Medical Education.*, 24, 507-511.

DOWIE, J. (1993) Clinical decision analysis: Background and introduction. IN LLEWELYN, H. & HOPKINS, A. (Eds.) *Analysing how we reach clinical decisions*. London: Royal college of physicians.

DOWIE, J. & ELSTEIN, A. (1988) *Professional judgment: A reader in clinical decision making.* , New York, Cambridge University Press.

EDWARDS, A., ELWYN, G., WOOD, F., ATWELL, C., PRIOR, L. and HOUSTON, H. (2004) Shared decision making and risk communication in practice: A qualitative study of GPs' experiences. *British Journal of General Practice*, 55, 6-13.

ELSTEIN, A. (1981) Educational programs in medical decision making. *Medical Decision Making*, 1, 70-73.

——— (1999) Heuristics and Biases: Selected Errors in Clinical Reasoning. *Academic Medicine*, 74, 791-794.

ELSTEIN, A., DAWSON-SAUNDERS, B. and BELZER, L. (1985) Instruction in Medical Decision Making: a report of two surveys. *Medical Decision Making*, 5, 229-232.

ELSTEIN, A., ROVNER, D. and ROTHERT, M. (1982) A preclinical course in decision making. *Medical Decision Making*, 2, 209-216.

ELSTEIN, A. & SCHWARTZ, A. (2002) Clinical problem solving and diagnostic decision making: selective review of the cognitive literature. *British Medical Journal*, 324, 729-732.

ELSTEIN, A. S., KAGAN, N., SHULMAN, L. S., JASON, H. and LOUPE, M. J. (1972) Methods and Theory in the Study of Medical Inquiry. *Journal of Medical Education*, 47, 85-92.

ELSTEIN, A. S., SHULMAN, L. S. and SPRAFKA, S. A. (1978) *Medical problem solving – an analysis of clinical reasoning*, Cambridge, Massachusetts: Harvard University Press.

EPPI (2007) Review guidelines for extracting data and quality assessing primary studies on educational research.

ERICSSON, K. (2007) An expert-performance perspective of research on medical expertise: the study of clinical performance. *Medical Education*, 41, 1124–1130.

ESSEX, B. & HEALY, M. (1994) Evaluation of a rule base for decision making in general practice. *British Journal of General Practice*, 44, 211-213.

EVA, K., NEVILLE, A. and NORMAN, G. (1998) Exploring the etiology of content specificity: factors influencing analogic transfer and problem solving *Academic Medicine* 73, S1-5.

EVA, K. W., HATALA, R. M., LEBLANC, V. R. and BROOKS, L. R. (2007) Teaching from the clinical reasoning literature: combined reasoning strategies help novice diagnosticians overcome misleading information. *Medical Education*, 41, 1152-1158.

EVA, K. W. & NORMAN, G. R. (2005) Heuristics and biases – a biased perspective on clinical reasoning. *Medical Education*, 39, 870.

EYSENCK, M. & KEANE, M. (2000) *Cognitive Psychology*, 4th ed ed., Psychology Press

FAMULARO, G., SALVINI, P., TERRANOVA, A. and GERACE, C. (2000) Clinical Errors in Emergency Medicine Experience at the Emergency Department of an Italian Teaching Hospital. *Academic Emergency Medicine*, 7, 1278.

FARNAN, J., JOHNSON, J., MELTZER, D., HUMPHREY, H. and ARORA, V. (2008) Resident uncertainty in clinical decision making and impact on patient care: a qualitative study *Quality and safety in healthcare*, 17, 122-126.

FAURIEL, I., MOUTEL, G., DUCHANGE, N., MONTUCLARD, L., MOUTARD, M., PIERRE COCHAT, C. and HERVE, C. (2005) Decision making concerning life-sustaining treatment in paediatric nephrology: professionals' experiences and values. *Nephrology dialysis transplantation*, 20, 2746–2750.

FISCHHOFF, B. (2002) Heuristics and biases in application. IN GILOVICH T, GRIFFIN D & KAHNEMAN D (Eds.) *Heuristics and biases: the psychology of intuitive judgment (Ch 41)*. Cambridge: Cambridge University Press.

FISCHHOFF, R. (1975) Hindsight is not equal to foresight: The effect of outcome knowledge on judgement under uncertainty. *Journal of Experimental Psychology. Human Perception & Performance.*, 1, 288-299.

FONTANA, A. & FREY, J. (2005) The interview: from neutral stance to political involvement. IN DENZIN, N. & LINCOLN, Y. (Eds.) *The Sage handbook of qualitative research.* Sage Publications, London.

FRIEDMAN, C., ELSTEIN, A., WOLF, F., MURPHY, G., FRANZ, T., HECKERLING, P., FINE, P., MILLER, T. and ABRAHAM, V. (1999) Enhancement of clinicians' diagnostic reasoning by computer-based consultation. *Journal of American Medical Association*, 282, 1851-1856.

FRIJLING, B., LOBO, C., HULSCHER, M., AKKERMANS, R., VAN DRENTH, B., PRINS, A., VAN DER WOUDEN, J. and GROL, R. (2003) Intensive support to improve clinical decision making in cardiovascular care: a randomised controlled trial in general practice. *Quality and Safety in Health Care*, 12, 181-187.

FRIJLING, B. D., LOBO, C., HULSCHER, M., AKKERMANS, R., BRASPENNING, J., PRINS, A., VAN DER WOUDEN, J. and GROL, R. (2002) Multifaceted support to improve clinical decision making in diabetes care: a randomized controlled trial in general practice. *Diabetic Medicine*, 19, 836-842.

FRISCH, D. & CLEMEN, R. T. (1994) Beyond Expectetd Utility: Rethinking Behavioral Decision Research. *Pschological Bulletin*, 116, 46-54.

GICK, M. & HOLLYOAK, K. (1983) Schema induction and analogical transfer. *Cognitve Psychology*, 15, 1-38.

GIFFORD, D., MITTMAN, B., FINK, A., LANTO, A., LEE, M. and VICKREY, B. (1996) Can a specialty society educate its members to think differently about clinical decisions? Results of a randomized trial. *Journal of General Internal Medicine*, 11, 664-672.

GLICK, T. H., WORKMAN, T. P. and GAUFBERG, S. V. (2000) Suspected Conversion Disorder Foreseeable Risks and Avoidable Errors. *Academic Emergency Medicine*, 7, 1272.

GMC (2009) Tomorrow's Doctors. London.

GODOLPHIN, W. (2003) The role of risk communication in shared decision making. *British Medical Journal*, 327, 692.

GOLDMAN, L., SAYSON, R., ROBBINS, S., COHN, LH., BETTMANN, M. and WEISBERG, M. (1983) The value of the autopsy in three different eras. *New England Journal of Medicine* 308, 1000-1005.

GRABER, M. (2003) Metacognitive training to reduce diagnostic errors: ready for prime time? . *Academic Medicine*, 78, 781.

——— (2009) Educational strategies to reduce diagnostic error: can you teach this stuff? *Advances in Health Sciences Education*, 14, 63-69.

GRABER, M., FRANKLIN, N. and GORDON, R. (2005) Diagnostic Error in Internal Medicine. *Archives of Internal Medicine* 165, 1493-1499.

GRABER, M., GORDON, R. and FRANKLIN, N. (2002) Reducing diagnostic errors in medicine: what's the goal? . *Academic Medicine*, 77, 981-992.

GRANT, D., KEIM, S. and TELFER, J. (2006) Teaching Bayesian Analysis to Emergency Medicine Residents. *The Journal of Emergency Medicine*, 31, 437-440.

GRANT, J. & DOWELL, J. (2002) A qualitative study of why general practitioners admit to community hospitals. *British Journal of General Practice*, 52, 628-634.

GRECO, P. & EISENBERG, J. (1993) Changing physicians' practices. *New England Journal of Medicine*, 329, 1271–1274.

GREENHALGH, T. (2002) Intuition and evidence – uneasy bedfellows? *British Journal of General Practice*, 52, 395-400.

GRIBBONS, B. & HERMAN, J. (1997) True and quasi-experimental designs. *Practical Assessment, Research & Evaluation*, 5.

GROEN, G. & PATEL, V. (1985) Medical problem solving: Some questionable assumptions. *Medical Education*, 19, 95-100.

GROOPMAN, J. (2007) *How doctors think*, Boston, Houghton Mifflin.

GRUPPEN, L., MARGOLIN, J., WISDOM, K. and GRUM, C. (1994) Outcome bias and cognitive dissonance in evaluating treatment decisions. *Academic Medicine*, 69, S57–S59.

GUEST, G., BUNCE, A. and JOHNSON, L. (2006) How Many Interviews Are Enough? An Experiment with Data Saturation and Variability. *Field Methods*, 18, 59-82.

GWEE, M. C. E. (2003) Medical Education: Issues, Trends, Challenges & Opportunities. *SMA News*, 35.

HAGGERTY, J., TUDIVER, F., BROWN, J., HERBERT, C., CIAMPI, A. and GUIBERT, R. (2005) Patients' anxiety and expectations: How they influence family physicians' decisions to order cancer screening tests. *Canadian Family Physician*, 51, 1658-1659.

HALL, J., HORGAN, T., STEIN, T. and ROTER, D. (2002) Liking in the Physician-Patient Relationship. *Patient Education and Counseling*, 48, 69-77.

HALL, J., MILBURN, M., ROTER, D. and DALTROY, L. (1998) Why Are Sicker Patients Less Satisfied With Their Medical Care? Tests of Two Explanatory Models. *Health Psychology* 17, 70-75.

HAMMOND, K. (2007) *Beyond rationality: the search for wisdom in a troubled time*, New York: Oxford University Press.

HAMMOND, K. R., HAMM, R. M., GRASSIA, J. and PEARSON, T. (1987) Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. *IEEE Transactions on Systems, Man, & Cybernetics*, 17, 753-770.

HARDEN, R. (2000) The integration ladder: a tool for curriculum planning and evaluation. *Medical Education*, 34, 551-557.

HARDEN, R., SOWDEN, S. and DUNN, W. (1984) Some educational strategies in curriculum development: The SPICES model. *Medical Education*, 18, 284-297.

HASSAN, N., ABDULLA, A., BAKATHIR, H., AL-AMOODI, A., AKLAN, A. and DE VRIES, T. (2000) The impact of problem-based pharmacotherapy training on the competence of rational prescribing of Yemen undergraduate students. *European Journal of Clinical Pharmacology*, 55, 873-876.

HATALA, R., NORMAN, G. and BROOKS, L. (2003) Practice makes perfect. the critical role of deliberate practice in the acquisition of ECG interpretation skills. *Advances in Health Science Education*, 8, 17-26.

HEDRICK, T. & YOUNG, J. (2008) The use of "war games" to enhance high-risk clinical decision-making in students and residents. *The American Journal of Surgery*, 195, 843-849.

HERSHBERGER, P., PART, H., MARKERT, R., COHEN, S. and FINGER, W. (1995) Teaching awareness of cognitive bias in medical decision making. *Academic Medicine*, 70.

HERSHBERGER, P. J., PART, H. M., MARKERT, R. J., COHEN, S. M. and FINGER, W. W. (1994) Development of a test of cognitive bias in medical decision making. *Academic Medicine*, 69, 839-841.

HIGGS, J. & JONES, M. (Eds.) (1995) *Clinical Reasoning in the Health Professions*, Oxford, Oxford : Butterworth-Heinermann.

HOFFRAGE, U. & GIGERENZER, G. (1998) Using natural frequencies to improve diagnostic inferences. *Academic Medicine*, 73, 538-540.

HOWE, K., HOLMES, M. and ELSTEIN, A. (1984) Teaching clinical decision making. *Journal of medical philosophy*, 9, 215-288.

HUNT, D., HAYNES, B., HANNA, S. and SMITH, K. (1998) Effects of Computer-Based Clinical Decision Support Systems on Physician Performance and Patient Outcomes: A Systematic Review. *Journal of the American Medical Association*, 280, 1339-1346.

JACKLIN, R., SEVDALIS, N., DARZI, D. and VINCENT, C. (2008) Mapping surgical practice decision making: an interview study to evaluate decisions in surgical care. *The American Journal of Surgery*, 195, 689-696.

JENSEN, G., SHEPARD, K. and HACK, L. (1990) The novice versus the experienced clinician: insights into the work of the physical therapist. *Physical Therapy*, 70 314-323.

JOSEPH, G. & PATEL, M. (1990) Domain knowledge and hypothesis generation in diagnostic reasoning. *Medical Decision Making*, 10, 31-46.

JUNGHANS, C., FEDER, G., TIMMIS, A., ELDRIDGE, S., SEKHRI, N., BLACK, N., SHEKELLE, P. and HEMINGWAY, H. (2007) Effect of patient-specific ratings vs conventional guidelines on investigation decisions in angina: Appropriateness of Referral and Investigation in Angina (ARIA) Trial. *Archives of Internal Medicine*, 167, 195-202.

KAHNEMAN, D., SLOVIC, P. and TVERSKY, A. E. (1982) *Judgment under uncertainty: heuristics and biases*, Cambridge University Press.

KAHNEMAN, D. & TVERSKY, A. (1972) Subjective probability: a judgment of representativeness. *Cognitive Psychology*, 3, 430-454.

——— (1973) On the psychology of prediction *Psychological review*, 80, 237-251.

KANTOWITZ, B., ROEDIGER, H. and ELMES, D. (2005) *Experimental psychology: understanding psychological research*, St Paul, MN, US, West Publishing Co.

KARAALP, A., AKICI, A., KOCABAS, E., LU, O. and OKTAY, S. (2003) What do graduates think about a two-week rational pharmacotherapy course in the fifth year of medical education? *Medical Teacher*, 25, 515-521.

KASSIRER, J. (1989) Diagnostic Reasoning. *Annals of Internal Medicine*, 110, 893-900.

——— (2010) Teaching Clinical Reasoning: Case-Based and Coached. *Academic Medicine* 85, 1118-1124.

KASSIRER, J. & KOPELMAN, R. (1991) *Learning Clinical Reasoning*, Baltimore: Williams and Wilkins.

KAWAMOTO, K., HOULIHAN, C. A., BALAS, E. A. and LOBACH, D. F. (2005) Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *British Medical Journal*, 330, 765-772.

KING, J. (1999) Giving feedback. *British Medical Journal*, 318, 2.

KING, N. (1994) The qualitative research interview. IN CASSELL, C. & SYMON, G. (Eds.) *Qualitative methods in organisational research* SAGE, London.

KIRCH, W. & SCHAFII, C. (1996) Misdiagnosis at a university hospital in 4 medical eras. *Medicine*, 75, 29-40.

KIRKPATRICK, D. (1998) *Evaluating training programs: the four levels* Berrett-Koehler publishers.

KOPP, V., STARK, R. and FISCHER, M. (2008) Fostering diagnostic knowledge through computer-supported, case-based worked examples: effects of erroneous examples and feedback. *Medical Education*, 42, 823-829.

KREBS, E., GARRETT, J. and KONRAD, T. (2006) The difficult doctor? Characteristics of physicians who report frustration with patients: an analysis of survey data. *BMC Health Services Research*, 6, 128.

KUHN, G. (2002) Diagnostic errors. *Academic Emergency Medicine* 9, 740-750.

KULATUNGA-MORUZI, C., BROOKS, L. R. and NORMAN, G. R. (2004) Using Comprehensive Feature Lists to Bias Medical Diagnosis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 30, 563-572.

KULIKOWSKI, C. A. (1970) Pattern Recognition Approach to Medical Diagnosis. *IEEE Transactions on Systems, Science and Cybernetics*, 6, 173-178.

KUMTA, S., TSANG, P., HUNG, L. and CHENG, J. (2003) Fostering critical thinking skills through a web-based tutorial programme for final year medical students. A randomized controlled study. *Journal of Educational Multimedia and Hypermedia*, 12, 267-273.

KURZENHAUSER, S. & HOFFRAGE, U. (2002) Teaching Bayesian reasoning: an evaluation of a classroom tutorial for medical students. *Medical Teacher*, 24, 516-521.

LAGERLOV, P., LOEB, M., ANDREW, M. and HJORTDAHL, P. (2000) Improving doctors' prescribing behaviour through reflection on guidelines and prescription feedback: a randomised controlled study. *Quality in Health Care*, 159-165.

LE GRAND, A., HOGERZEIL, H. and HAIJJER-RUSKAMP, F. (1999) Intervention research in rational use of drugs: a review. *Health and policy planning*, 14, 89-102.

LEE, A., JOYNT, G. M., HO, A. M. H., GIN, T. and HAZLETT, C. B. (2007) Effect of an integrated teaching intervention on clinical decision analysis: a randomized, controlled study of undergraduate medical students. *Medical Teacher*, 29, 231-236.

LÉGARÉ, F., RATTÉ, S., STACEY, D., KRYWORUCHKO, J., GRAVEL, K., GRAHAM, I. and TURCOTTE, S. (2010) Interventions for improving the adoption of shared decision making by healthcare professionals (Review). *The Cochrane Library*.

LERNAU, O. (1989) Problem-solving instruction during the clinical clerkship: Description and preliminary evaluation of a programme. *Medical Education*, 23, 179-183.

LINCOLN, M., TURNER, C., HAUG, P., WARNER, H., WILLIAMSON, J., BOUHADDOU, O., JESSEN, S., SORENSON, D., CUNDICK, R. and GRANT, M. (1991) Iliad training enhances medical students' diagnostic skills. *Journal of Medical Systems*, 15, 93-110.

LOCKEY, A. & HARDERN, R. (2001) Decision making by emergency physicians when assessing cardiac arrest patients on arrival at hospital. *Resucitation*, 50, 51-56.

LUNDBORG, C., WAHLSTROM, R., DIWAN, V., OKE, T., MARTENSON, D. and TOMSON, G. (1999) Combining feedback from simulated cases and prescribing. Design and implementation of an educational intervention in primary care in Sweden. *International Journal of Technology Assessment in Health Care*, 15, 458-472.

MARGOLIS, C., BARLOON S and N., B. (1982) A required course in decision-making for preclinical medical students. *Journal of Medical Education*, 57, 184-190.

MATHERS, N., JONES, N. and HANNAY, D. (1995) Heartsink patients: a study of their general practitioners. *British Journal of General Practice*, 45, 293-296.

MAY, T. (2002) *Qualitative research in action*, London: SAGE.

MAYOU, R. (1978) Psychiatric decision-making by medical students. *British Journal Psychiatry*, 132, 191-194.

MCKIMM, J. (2010) Current trends in undergraduate medical education: Program and Curriculum Design. *Samoa Medical Journal*, 2, 40-48.

MCKINLAY, J., POTTER, D. and FELDMAN, H. (1996) Non-medical influences on medical decision-making. *Social Science & Medicine*, 42, 769-776.

MEANS, B., SALAS, E., CRANDALL, B. and JACOBS, J. O. (1993) Training Decision Makers for the Real World in. IN KLEIN, G. A., ORASANU, J., CALDERWOOD, R. & ZSAMBOK, C. E. (Eds.) *Decision Making in Action: Models and Methods* Ablex Publishing, Corporation Norwood, New Jersey.

METCALFE, J. & SHIMAMURA, A. (Eds.) (1994) *Metacogition: knowing about knowing* Cambridge, London, The MIT Press.

MILLER,G. E. (1990) The assessment of clinical skills/competence/performance. *Academic Medicine*, 65, S63-S67.

MOREIRA, J., BISOFFI, Z., NARVAEZ, A. and VAN DEN ENDE, J. (2008) Bayesian clinical reasoning: does intuitive estimation of likelihood ratios on an ordinal scale outperform estimation of sensitivities and specificities? *Journal of Evaluation in Clinical Practice*, 14, 934-940.

MORSE, J. (1995) The significance of saturation. *Qualitative Health Research*, 5, 147–149.

MRC (2008) Developing and evaluating complex interventions: new guidance.

MURRAY, T., CUPPLES, R., BARBERM, J., DUNNM, W., SCOTTM, D. and HANNAY, D. (1977) Teaching decision making to medical undergraduates by computer assisted learning. *Medical Education*, 11, 262-264.

MYLOPOULOS, M. & WOODS, N. (2009) Having our cake and eating it too: seeking the best of both worlds in expertise research. *Medical Education*, 43, 406-413.

NAKAMURA, C. (2008) The effects of specific support to hypothesis generation on the diagnostic performance of medical students. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 68, 4264.

NENDAZ, M. & BORDAGE, G. (2002) Promoting diagnostic problem representation. *Medica Education*, 36, 761–767.

NEWELL, A. & SIMON, H. (1972) *Human problem solving*, Englewood Cliffs, NJ: Prentice-Hall.

NEWTON-SYMS, F., DAWSON, P., COOKE, J., FEELY, M., BOOTH, T., JERWOOD, D. and CALVERT, R. (1992) The influence of an academic representative on prescribing by general practitioners. *British Journal of Clinical Pharmacology*, 33, 69-73.

NILSSON, G., HJEMDAHL, P., HASSLER, A., VITOLS, S., WALLEN, N. and KRAKAU, I. (2001) Feedback on prescribing rate combined with problem-orientated pharmacotherapy education as a model to improve prescribing behaviour among general practitioners. *European Journal of Clinical Pharmacology*, 56, 843-848.

NOE, R. & SCHMITT, N. (1986) The influence of trainee attitudes on training effectiveness: test of a model. *Personnel Psychology* 39, 497-523.

NOGUCHI, Y., MATSUI, K., IMURA, H., KIYOTA, M. and FUKUI, T. (2004) A Traditionally Administered Short Course Failed to Improve Medical Students' Diagnostic Performance. *Journal of General Internal Medicine*, 19, 427-432.

NORMAN, G. (2005) Research in clinical reasoning: past history and current trends. *Medical Education*, 39, 418-427.

——— (2009) Dual processing and diagnostic errors. *Advances in Health Sciences Education*, 14, 37-49.

NORMAN, G., COBLENTZ, C., BROOKS, L. and BABCOOK, C. (1992) Expertise in visual diagnosis: a review of the literature. *Academic Medicine*, 67, S78-83.

NORMAN, G. & EVA, K. (2010) Diagnostic error and clinical reasoning. *Medical Education*, 44, 94-100.

NORMAN, G., EVA, K. and SCHMIDT, H. (2005) Implications of psychology-type theories for full curriculum interventions. *Medical Education*, 39, 243-249.

NORMAN, G., ROSENTHAL, D., BROOKS, L., ALLEN, S. and MUZZIN, L. (1989) The Development of Expertise in Dermatology. *Archives of Dermatology* 125, 1063-1068.

NORMAN, G., TUGWELL, P., FEIGHTNER, J., MUZZIN, L. and JACOBY, L. (1995) Knowledge and clinical problem solving. *Medical Education*, 19, 344-356.

O'CONNELL, D., HENRY, D. and TOMLINS, R. (1999) Randomised controlled trial of effect of feedback general practitioners' prescribing in Australia. *British Medical Journal*, 318, 507-511.

O'CONNOR, A., ROSTOM, A., FISET, V., TETROE, J., ENTWISTLE, V., LLEWELLYN-THOMAS, H., HOLMES-ROVNER, M., BARRY, M. and JONES, J. (1999) Decision aids for patients facing health treatment or screening decisions: a systematic review. *British Medical Journal*, 319, 731-734.

PATEL, V. & GROEN, G. (1986) Knowledge based solution strategies in medical reasoning. *Cognitive Science*, 10, 91-116.

PAUKER, S. & PAUKER, S. (1999) What is a good decision? *Effective Clinical Practice* 2, 194-196.

PAYNE, J. & BETTMAN, J. (2004) Walking with the scarecrow: the information-processing approach to decision research (Ch 6). IN DJ, I. K. & EDITORS, H. N. (Eds.) *Blackwell handbook of judgment and decision making.* Oxford: Blackwell Publishing Ltd.

PEABODY, J., LUCK, J., GLASSMAN, P., DRESSELHAUS, T. and LEE, M. (2000) Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *Journal of American Medical Association*, 283, 1715-1722.

PEABODY, J., LUCK, J., GLASSMAN, P., JAIN, S., HANSEN, J., SPELL, M. and LEE, M. (2004) Measuring the Quality of Physician Practice by Using Clinical Vignettes: A prospective Validation Study. *Annals of Internal Medicine*, 141, 771-780.

PETTICREW, M. & ROBERTS, H. (2006) *Systematic reviews in the social sciences: a practical guide*, Oxford, Blackwell.

PRETZ, J. (2008) Intuition versus analysis: Strategy and experience in complex everyday problem solving. *Memory & Cognition*, 36, 554-566.

RACE, P. (2005) *Making learning happen: a guide for post-compulsory education*, Sage Publications.

RELLO, J., LORENTE, C., BODÍ, M., DIAZ, E., RICART, M. and KOLLEF, M. (2002) Why Do Physicians Not Follow Evidence-Based Guidelines for Preventing Ventilator-Associated Pneumonia? A Survey Based on the Opinions of an International Panel of Intensivists. *Chest*, 122, 656-661.

RITCHIE, J. & LEWIS, J. (2003) *Qualitative Research Practice: A guide for social science students and researchers*, London: SAGE.

ROBSON, C. (2002) *Real World Research*, 2nd ed. Oxford, Blackwell Publishing.

ROGERS, J., SWEE, D. and ULLIAN, J. (1991) Teaching medical decision making and students' clinical problem solving skills. *Medical Teacher*, 13, 157-164.

ROGERS, P., GRENVIK, A. and WILLENKIN, R. (1995) Teaching medical students complex cognitive skills in the intensive care unit. *Critlcal Care Medicine*, 23, 1933-1935.

ROSS, B. (1987) This is like that: the use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 629-639.

ROSS, B. & KENNEDY, P. (1990) Generalizing From the Use of Earlier Examples in Problem Solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16, 42-55.

ROUND, A. P. (1999) Teaching clinical reasoning - a preliminary controlled study. *Medical Education*, 33, 480-483.

RUIZ, J., MINTZER, M. and LEIPZIG, R. (2006) The Impact of E-Learning in Medical Education. *Academic Medicine*, 81, 207-212.

SCHMIDT, H., NORMAN, G. and BOSHUIZEN, H. (1990) A cognitive perspective on medical expertise: theory and implications. *Academic Medicine*, 65, 611-621.

SCHMIDT, H. G. & RIKERS, R. M. J. P. (2007) How expertise develops in medicine: knowledge encapsulation and illness script formation. *Medical Education*, 41, 1133-1139.

SCHOLDER ELLEN, P. (1994) Do we know what we need to know? Objective and subjective knowledge effects on pro-ecological behaviors. *Journal of Business Research*, 30, 43-52.

SCHUTT, R. (2006) *Investigating the social world: the process and practice of research*, 5th ed., Thousand Oaks, Calif.: SAGE Publications.

SCHUWIRTH, L. (2002) Can clinical reasoning be taught or can it only be learned? *Medical Education*, 36, 695-696.

SCHWARTZ, A. & BERGUS, G. (2008) *Medical Decision Making: a physician's guide*, Cambridge: Cambridge University Press.

SCHWARTZ, R., DONNELLY, M., NASH, P., JOHNSON, S., YOUNG, B. and GRIFFEN, W. (1992) Problem-Based Learning: An Effective Educational Method for a Surgery Clerkship. *Journal of Surgical Research*, 53, 326-330.

SEDGWICK, P. & HALL, A. (2003) Teaching medical students and doctors how to communicate risk: Combining the teaching of statistics with communication skills. *British Medical Journal*, 327, 694-695.

SERVAIS, E., LAMORTE, W., AGARWAL, S., MOSCHETTI, W., MALLIPATTU, S. and MOULTON, S. (2006) Teaching Surgical Decision-Making: An Interactive, Web-Based Approach. *Journal of Surgical Research* 134, 102-106.

SHANTEAU, J. (1992) How much information does an expert use? Is it relevant? *Acta Psychologica*, 81, 75-86.

SHEKELLE, P., KRAVITZ, R., BEART, J., MARGER, M., WANG, M. and LEE, M. (2000) Are nonspecific practice guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on physician decision making. *Health Services Research*, 34, 1429-1448.

SHEPPERD, J. & KOCH, E. (2005) Pitfalls in Teaching Judgment Heuristics. *Teaching of Psychology*, 32, 43-46.

SHOJANIA, K., BURTON, E., MCDONALD, K. and GOLDMAN, L. (2003) Changes in rates of autopsydetected diagnostic errors over time. *Journal of American Medical Association*, 289, 2849-2856.

SIMON, H. (1972) Theories of bounded rationality, Ch 8 in. IN MCGUIRE, C. & RADNER, R. E. (Eds.) *Decision and Organization*. Amsterdam: North-Holland Publishing Company.

SINTCHENKO, V., COIERA, E., IREDELL, J. and GILBERT, G. (2004) Comparative impact of guidelines, clinical data, and decision support on prescribing decisions: an interactive web experiment with simulated cases. *Journal of the American Medical Informatics Association*, 11, 71-77.

SMITH, J. (2008) *Qualitative psychology : a practical guide to research methods*, 2nd ed., Los Angeles, Calif, London: SAGE.

SOX, H. (1999) What makes a good decision. *Effective Clinical Practice*, 2, 196-197.

SPENCER, J. & JORDAN, R. (1999) Learner centred approaches in medical education. *British Medical Journal* 318, 1280-1283.

STANOVICH, K. & WEST, R. (2000) Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, 23, 645-665.

TABACHNICK, B. & FIDELL, L. (1996) *Using multivariate statistics*, 3 ed., New York: HarperCollins.

TAMBLYN, R., HUANG, A., PERREAULT, R., JACQUES JAMES, A., HANLEY, D., MCLEOD, P. and LAPRISE, R. (2003) The medical office of the 21st century (MOXXI): effectiveness of computerized decision-making support in reducing inappropriate prescribing in primary care. . *Canadian Medical Association Journal*, 169, 549-556.

TAUT, S. & ALKIN, M. (2003) Program Staff Perceptions of Barriers to Evaluation Implementation. *American Journal of Evaluation*, 24, 213-226.

TERRELL, K. M., PERKINS, A. J., DEXTER, P. R., HUI, S. L., CALLAHAN, C. M. and MILLER, D. K. (2009) Computerized Decision Support to Reduce Potentially Inappropriate Prescribing to Older Emergency Department Patients: A Randomized, Controlled Trial. *Journal of the American Geriatrics Society* 57, 1388–1394.

THOMAS, A., BOXALL, E., LAHA, S., DAY, A. and GRUNDY, D. (2008) An educational and audit tool to reduce prescribing error in intensive care. *Quality and Safety in Health Care*, 17, 360-363.

THOMPSON, C. & DOWDING, D. (2002) *Clinical decision making and judgement in nursing*, Edinburgh, Churchill Livingstone.

TORGERSON, C. (2003) *Systematic Reviews*, London, Continuum.

TVERSKY, A. & KAHNEMAN, D. (1973) Availability: a heuristic for judging frequency and probability *Cognitive Psychology*, 5, 207-232.

——— (1974) Judgment under uncertainty: heuristics and biases. *Science*, 185, 1124-1131.

VAN DER VELDE, G. (2005) Clinical decision analysis: an alternate, rigorous approach to making clinical decisions and developing treatment recommendations. *Journal of the Canadian Chiropractic Association*, 49, 258-263.

VENINGA, C., DENIG, P., ZWAAGSTRA, R. and HAAIJER-RUSKAMP, F. (2000) Improving drug treatment in general practice. *Journal of Clinical Epidemiology*, 53, 762-772.

VENINGA, C., LAGERLØV, P., WAHLSTRÖM, R., MUSKOVA, M., DENIG, P., BERKHOF, J., KOCHEN, M. and HAAIJER-RUSKAMP, F. (1999) Evaluating an Educational Intervention to Improve the Treatment of Asthma in Four European Countries. *American Journal of Respiratory and Critical Care Medicine*, 160, 1254-1262.

VOLLEBREGT, J., METZ, J., DE HAAN, M., RICHIR, M., HUGTENBURG, J. and DE VRIES, T. (2005) Curriculum development in pharmacotherapy: testing the ability of preclinical medical students to learn therapeutic problem solving in a randomized controlled trial. *Br J Clin Pharmacol*, 61, 345-351.

WARNER, H., WOOLLEY, F. and KANE, R. (1974) Computer assisted instruction for teaching clinical decision making. *Computers and Biomedical Research*, 7, 564-574.

WATSON, E., CLEMENTS, A., YUDKIN, P., ROSE, P., BUKACH, C., MACKAY, J., LUCASSEN, A. and AUSTOKER, J. (2001) Evaluation of the impact of two educational interventions on GP management of familial breast/ovarian cancer cases: a cluster randomised controlled trial. *British Journal of General Practice*, 51, 817-821.

WELSH, M., BEGG, S. and BRATVOLD, R. (2007) Efficacy of bias awareness in debiasing oil and gas judgments. *Proceedings of the 29th Annual Conference of the Cognitive Science Society*. Nashville, TN, USA.

WIGTON, R., PATIL, K. and HOELLERICH, V. (1986) The effect of feedback in learning clinical diagnosis. *Journal of Medical Education*, 16, 816-822.

WIGTON, R., POSES, R., COLLINS, M. and CEBUL, R. (1990) Teaching old dogs new tricks: Using cognitive feedback to improve physicians' diagnostic judgments on simulated cases. *Academic Medicine*, 65, S5-S6.

WILLIG, C. (2001) *Introducing qualitative research in psychology: Adventures in theory and method* Buckingham, Open University Press.

WILSON, R., RUNCIMAN, W., GIBBERD, R., HARRISON, B., NEWBY, L. and HAMILTON, J. (1995) The Quality in Australian Health Care Study. *The Medical Journal of Australia*, 163, 458-471.

WILSON VANVOORHIS,C. & MORGAN, B. (2007). Understanding Power and Rules of Thumb for Determining Sample Size. *Tutorials in Quantitative Methods for Psychology* 3, 43-50.

WINDISH, D., PRICE, E., CLEVER, S., MAGAZINER, J. and THOMAS, P. (2005) Teaching medical students the important connection between communication and clinical reasoning. *Journal of General Internal Medicine*, 20, 1108-1113.

WOLF, F., GRUPPEN, L. and BILLI, J. (1988) Use of the competing hypothesis heuristic to reduce 'pseudodiagnosticity'. *Journal of Medical Education*, 63, 548-554.

WOODS, N. (2007) Science is fundamental: the role of biomedical knowledge in clinical reasoning. *Medical Education*, 41, 1173–1177.

WU, G., ZHANG, J. and GONZALEZ, R. (2004) Decision under risk. IN KOEHLER, D. & HARVEY, N. (Eds.) *Blackwell handbook of judgement and decision making.* Blackwell publishing.

WUTOH, R., BOREN, S. A. and BALAS, E. A. (2004) eLearning: A review of Internet-based continuing medical education. *Journal of Continuing Education in the Health Professions*, 24, 20.

YATES, J. F., VEINOTT, E. S. and PATALANO, A. L. (2003) Hard decisions, bad decisions: On decision quality and decision aiding. IN SCHNEIDER, S. & SHANTEAU, J. (Eds.) *Emerging perspectives on judgment and decision research (13-63).* New York: Cambridge University.

8. Appendices

8.1. Systematic review search strategy

8.1.1. OVID

- Searched Medline, Embase and PsycInfo separately so relevant articles were not missed.
- Searched from 1950 to May 2010.
- Three topic headings developed and recorded as many different similar words.

Topic 1 – terms related to 'intervention to facilitate'	Topic 2 – terms related to 'doctors'	Topic 3 – terms related to 'clinical decision making'
Intervention	Doctors	Decision making
Teach	Physician	Clinical decision making
Train	Clinician	Clinician decision making
Educate	Medic	Physician decision making
Instruct	Medical Student	Medical decision making
Guide	Medical practitioner	
Learn	Trainee	Clinical judgement
Clerkship	Postgraduate	Clinician judgement
Internship	Undergraduate	Physician judgement
		Medical judgement
Facilitate	Consultant	
Aid	Registrar	Clinical reasoning
Enhance	House officer	Medical reasoning
Help	Specialist	Critical thinking
Assist	General practitioner	Problem solving
Develop	Intern	
Promote	Resident	
Evaluate		

- Using the 'map term to subject heading' function, inserted following subheadings:
  - Topic 1: physicians/ or hospitalists/ or physicians, family/ or physicians, women/

- Topic 2: Students, Medical
- Topic 3: Consultants
- Topic 4: medical education

- Key words searches carried out. Unticked 'map term to subject heading'.
- Used truncation * for similar words like train* to find trainer, training etc.
- Used '?' symbol for multiple ways of spelling a word e.g. judg?ment
- When all of key word searches done, ticked all boxes and 'or' to combine all data sets.
- This resulted in 3 main data sets.
- Combined the 3 data sets into 1 final large data set to look through using 'and' function.
- Articles limited to English language only.
- Spot checked first few pages to check the suitability of search terms.

8.1.2. Other databases

Search features in other databases not as advanced as OVID, key words typed in like an equation in the following databases;

- Cochrane library (from earliest date to My 2010)
- Web of Science (from 1898-May 2010)
- Cambridge Scientific Abstracts (1900-May 2010)
- ERIC (from 1966-May 2010)

Key words inserted into each database as follows;

- intervention* or teach* or train* or educate* or guide* or clerkship or internship or facilitat* or aid* or enhanc*or develop* or evaluat* or medical education

and

- doctor* or physician* or clinician* or medic* or trainee* or postgraduate* or undergraduate* or general practitioner* or GP* or intern* or resident*

and

- decision making* or judg?ment* or clinical reasoning or medical reasoning or critical thinking or problem solving

8.2. Data extraction form

Study ID __________

Details of study

Author(s) ____________________

Title ____

____

Reference ____

____

Country of origin ________________

☐ Full paper ☐ Abstract/summary only

Aim of study

Fill in with authors words

Discussion of 'good' metacognitive skills (i.e. what makes good decision making)

☐ Yes ☐ No

Details

Theory used

- ☐ None
- ☐ Medical Education
- ☐ Economics
- ☐ Psychology
- ☐ Mathematics
- ☐ Other

Details ______________________________

____

Design

Fill in with authors words

Sampling method

- ☐ Convenience/opportunistic
- ☐ Random
- ☐ All class/year
- ☐ Other *(state)*
- ☐ Purposive
- ☐ Volunteer
- ☐ Not stated

Sample size

Numbers participated

☐ Total ☐ Intervention group

☐ Control group ☐ Attrition (participants lost)

Numbers participated to completion

☐ Intervention group ☐ Control group

*Other details*____

____

____

Other sample characteristics

select all that apply

☐ Medical undergraduates *Stage of training (state)* __________

☐ Doctors ☐ Other professionals

Position / speciality
(state) ____

____

Other details

Intervention

Type

- ☐ Metacognition (teaching how people think)
- ☐ Expert heuristic (decision aids- teaching PBL, guidelines, decision analysis)
- ☐ Other

Further details

Setting

☐ University	☐ Primary care
☐ Hospital	☐ Laboratory
☐ Not stated	☐ Other (state)

Integration of intervention

- ☐ Integrated into medical course
- ☐ Integrated into behavioural sciences course
- ☐ Isolated/one off
- ☐ Unclear

Participants' task

Measures

Patient management - specific

☐ Diagnosis ☐ Treatment plan/prescribing
☐ History taking ☐ Skills acquisition

Metacognitive skills - general

☐ Decision making ☐ Reasoning
☐ Judgement ☐ Problem solving
☐ Critical thinking ☐ *Other (details)*

Validity and reliability of measures

Measurement type (*e.g. questionnaire, process tracing, observation*)	Description *(e.g. diagnostic thinking inventory)*	Valid/established?

Timing of measurement *(select all that apply)*

☐ Before intervention only ☐ During intervention
☐ After intervention only ☐ Before and after intervention

Further comments

Results

Effective training? (As stated by authors)

☐ Yes → ☐ Significant
☐ Non-significant
☐ Mix of significant / non-significant
☐ Unclear
☐ N/A

☐ No

Summary of conclusions

Assessing quality of study

Yes = 2 Unclear = 1 No = 0 N/A = Omit from total score

Theory	**Yes**	**Unclear**	**No**	**N/A**
Referred to in introduction and discussion				
Operationlised				
Evaluated				

Scientific rigour	**Yes**	**Unclear**	**No**	**N/A**
Clear aims				
Aims are meaningful/justified				
Appropriate sample used				
Appropriate comparator (similar groups)				
Mostly validated/established measures				
Measures used at appropriate times				

Coherence of study	**Yes**	**Unclear**	**No**	**N/A**
Appropriate method for aims				
Appropriate measures				
Meaningful results				
Conclusion is consistent with results				

Generalisability	**Yes**	**Unclear**	**No**	**N/A**
Sample is representative of study population (i.e. represents who they claim intervention is aimed at)				
Limitations acknowledged				

Further details

Total quality score ___________

8.3. Ethical approval for questionnaire study

From: Darren Shickle
To: Vikram Jha
Cc: Hilary Bekker; Leila Mehdizadeh; Deborah Murdoch-Eaton
Date: 24 September 2009 15:25:11
Subject: RE: Ethics

Dear Vikram,

Normally I would say that this would fine and would treat it as though Deborah has taken chair's action on behalf of EdREC to say that ethics review was not required. However, given that this is for Leila's PhD, we would need to make sure that an examiner doesn't cause problems down the line. Thus if Leila would like extra reassurance, we could do a formal review if she wanted to complete the ethics application form. Alternatively I have re-read the protocol and questionnaire that you sent me, and I do not see anything ethically problematic. In particular, I have considered the following:

- The introduction to the questionnaire explains the purpose and together with the act of completion and return infers informed consent.
- The questionnaire is anonymous, although if returned by e-mail, could allow the questionnaire to be identified, but I assume that no identifying information will be transferred to the questionnaire.
- I am also assuming that members of the distribution list will have given implied consent to receive questionnaires in this way ie it happens on a regular basis.
- The subject of the questionnaire is not sensitive.

On this basis I am happy to take chair's action on behalf of the Medicine and Health Faculty Research Ethics Committee to approve Leila's research. She should retain this e-mail as proof of this decision, although of course she can still submit an ethics application if she wants additional evidence.

Regards

Darren

Professor Darren Shickle
Head of the Academic Unit of Public Health
Institute of Health Sciences
University of Leeds
Room G.30
Charles Thackrah Building
101 Clarendon Road
Leeds LS2 9LJ
Telephone: +44 113 343 7213
Fax: +44 113 343 6997
E-mail: d.shickle@leeds.ac.uk

8.4. Cover letter from supervisor for questionnaire study

From: Evidence Based Shared Decision Making [mailto:SHARED-L@LIST.MSU.EDU] On Behalf Of Hilary Bekker
Sent: Tuesday, October 27, 2009 12:49 PM
To: SHARED-L@LIST.MSU.EDU
Subject: for teachers of medical decision making

Dear All

Re: a very short questionnaire.

One of my PhD students, Leila Mehdizadeh, is carrying out a survey to find out what type of training doctors / medical students are receiving about medical decision making. The survey is building on Arthur Elstein's survey from 20 years ago.

If you have taught or are teaching doctors about some aspect of medical decision making, Leila would be grateful if you would complete the attached survey. If you know of a colleague who is involved, please forward the email to them.

Please respond directly to Leila's at psc3lm@leeds.ac.uk; Or Leila Mehdizadeh, room G.02, Leeds Institute of Health Sciences, Charles Thackrah Building, University of Leeds, 101 Clarendon Road, Leeds LS2 9LJ, United Kingdom

Thank you in anticipation of your help

Hilary

Dr Hilary L Bekker
Chartered Health Psychologist
Senior Lecturer in Behavioural Sciences
Leeds Institute of Health Sciences, Faculty of Medicine and Health, University of Leeds
Charles Thackrah Building - Room 1.10, 101 Clarendon Road, Leeds LS2 9 LJ
Tel: +44 (0) 113 343 2726
Email: h.l.bekker@leeds.ac.uk
http://www.leeds.ac.uk/lihs/psychiatry/staff/bekker.htm

8.5. Medical decision making questionnaire

TEACHING DOCTORS ABOUT MEDICAL DECISION MAKING: A BRIEF SURVEY

Please complete this survey if you have taught doctors about an aspect of medical decision making; either courses designed for doctors to help improve their decision making skills and/or those of their patients. As no gold standard has been suggested to guide educators' medical decision making programmes, I am asking for your experiences and views about the necessary component parts for a medical decision making course.

The survey is part of my PhD investigating courses designed to teach doctors about medical decision making. I am supervised by Dr Hilary Bekker (Leeds Institute of Health Sciences) and Drs Vikram Jha and Naomi Quinton (Leeds Institute of Medical Education). Thank you for your help, Leila Mehdizadeh (**psc3lm@leeds.ac.uk).**

1. For which organisations do you (or have) deliver a medical decision making course? (tick all that apply)

Organisation	**Please list their affiliation (Leeds Uni, hospital trust, RCOG, FIMDM, etc)**
University	
Medical Professional Body	
Conferences	
Public Health Facility	
Private Health Facility	
Other *(please state)*	

2. How have you delivered the medical decision making course(s)? (tick all that apply)

A half-day specialist session		A short 1-2 day specialist course	
Part of a module on a broader topic		As a module on decision making	
Other *(please state)*			

3. What is the focus of your medical decision making course(s)? (tick all that apply)

Helping doctors to make better clinical decisions		Helping doctors to develop some general awareness of decision making (meta cognition)	
Helping doctors to make better team decisions		Helping doctors' skills in patient-centred care or shared decision making	
Informing doctors about social sciences			
Other *(please state)*			

4. What theoretical perspectives do you draw on during your teaching of the course(s)? (tick all that apply)

Classical/ normative/ rational decision theories		Descriptive information processing theories	
Expert models of clinical decision making/naturalistic theories		Heuristics and biases/ framing	
Self regulation theories		Social cognition models	
None			
Other *(please state)*			

5. What applications do you refer to during your teaching of the course(s)? (tick all that apply):

Decision analysis		Other algorithms	
Conjoint analysis		Smart heuristics	
Patient decision aids		Utility elicitation methods	
Risk presentation		Bayes' Theorem	
Other (please list)			

6. Is the effectiveness of the decision making part of the course(s) explicitly evaluated? (tick one response)

Yes		No	

7. What do you think should be included in a medical decision making course? (tick all that apply)

Clinical reasoning strategies (e.g. hypothetico-deductive method, pattern recognition)		Heuristics and overcoming bias	
Risk Perception		Probabilities (e.g. likelihood ratios, pre/post test probabilities of disease)	
Diagnostic test accuracy (sensitivity, specificity)		Decision analysis	
Cost-effectiveness analysis		Clinical practice guidelines	
Decision aids			
Other (please list)			

8. What method is most useful for teaching medical decision making? (tick all that apply)

Paper based clinical scenarios		Computerised clinical scenarios	
High-fidelity simulation (manikins)		Standardised patients/actors	
Real patient care experience (e.g. clinics, ward rounds)		Erroneous clinical examples	
Anecdotes of doctors			
Other (please list)			

9. How much do you agree with the following statements about how and when to teach medical decision making? (tick one response for each statement)

	Not relevant	Possibly relevant	Very relevant
Integrated with social sciences courses in undergraduate training			
Integrated into communication skills courses in undergraduate training			
Integrated with clinical attachment in undergraduate training			
An optional specialist topic in undergraduate training			
Integrated into professional training post first degree			
An optional component in continuing professional development portfolios			
A work based learning activity			

10. How much do you agree with the following statements about their assessment and impact? (tick one response for each statement)

	Not relevant	Possibly relevant	Very relevant
Doctors' decision making should be assessed in the workforce			
Formal decision making training will increase patient safety			
It is unlikely formal decision making training will gain wide support in medical education			

11. What advice do you have about medical decision making training for other educators?

Thank you for taking the time to complete this questionnaire. Please return the completed questionnaire to the box at the conference reception desk or post by the 2nd November to; Leila Mehdizadeh room G.02 Leeds Institute of Health Sciences, Charles Thackrah Building, University of Leeds, 101 Clarendon Road, Leeds LS2 9LJ, United Kingdom

8.6. NHS ethical approval for interview study

NHS

National Research Ethics Service

Leeds (East) Research Ethics Committee
Room 5.2, Clinical Sciences Building
St James's University Hospital
Beckett Street
Leeds
LS9 7TF

Telephone: 0113 2065652
Facsimile: 0113 2066772

1 April 2009

Ms Leila Mehdizadeh
PhD Research Student
Leeds Institute of Health Sciences
Room G.02
Charles Thackrah Building
101 Clarendon Road
LS2 9LJ

Dear Ms Mehdizadeh

Full title of study:	**Clinicians' Views on Effective Clinical Decision Making: An Interview Study**
REC reference number:	**09/H1306/11**

Thank you for your letter of 25 March 2009, responding to the Committee's request for further information on the above research and submitting revised documentation.

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, subject to the conditions specified below.

Ethical review of research sites

The Committee has designated this study as exempt from site-specific assessment (SSA). The favourable opinion for the study applies to all sites involved in the research. There is no requirement for other Local Research Ethics Committees to be informed or SSA to be carried out at each site.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission at NHS sites ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Interview Schedules/Topic Guides	2	06 February 2009
Compensation Arrangements		02 October 2008
Letter from Sponsor		06 February 2009
Protocol	1	06 February 2009
Investigator CV		06 February 2009
Application		06 February 2009
CV for Dr Bekker		
Response to Request for Further Information		25 March 2009
Participant Consent Form	2.0	23 March 2009
Participant Information Sheet	4.0	23 March 2009

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

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "After ethical review –guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

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@nres.npsa.nhs.uk.

09/H1306/11	**Please quote this number on all correspondence**

This Research Ethics Committee is an advisory committee to Yorkshire and The Humber Strategic Health Authority

The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England.

With the Committee's best wishes for the success of this project

Yours sincerely

Dr John Holmes
Chair

Email: ann.tunley@leedsth.nhs.uk

Enclosures: "After ethical review – guidance for researchers"

Copy to: Ms Rachel De Souza, University of Leeds

R & D Department, Leeds Teaching Hospitals NHS Trust

This Research Ethics Committee is an advisory committee to Yorkshire and The Humber Strategic Health Authority
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England.

The Leeds Teaching Hospitals **NHS**
NHS Trust

01/06/2009

Research & Development
Leeds Teaching Hospitals NHS Trust
34 Hyde Terrace
Leeds
LS2 9LN

Tel: 0113 392 2878
Fax: 0113 392 6397

r&d@leedsth.nhs.uk
www.leedsteachinghospitals.com

Ms Leila Mehdizadeh
Leeds Institute of Health Sciences
Room G.02 Charles Thackrah Building
101 Clarendon Road
Leeds
LS2 9LJ

Dear Ms Leila Mehdizadeh

Re: LTHT R&D Approval of: Clinicians Views on Effective Clinical Decision Making : An Interview Study
LTHT R&D Number: UI09/8927
MREC: 09/H1306/11

I confirm that this study has R&D approval and the study may proceed at The Leeds Teaching Hospitals NHS Trust (LTHT). This organisational level approval is given based on the information provided in the documents listed below.

In undertaking this research you must comply with the requirements of the *Research Governance Framework for Health and Social Care* which is mandatory for all NHS employees. This document may be accessed on the R&D website http://www.leedsth.nhs.uk/sites/research_and_development/

R&D approval is given on the understanding that you comply with the requirements of the *Framework* as listed in the attached sheet "Conditions of Approval".

If you have any queries about this approval please do not hesitate to contact the R&D Department on telephone 0113 392 2878.

Indemnity Arrangements

The Leeds Teaching Hospitals NHS Trust participates in the NHS risk pooling scheme administered by the NHS Litigation Authority 'Clinical Negligence Scheme for NHS Trusts' for: (i) medical professional and/or medical malpractice liability; and (ii) general liability. NHS Indemnity for negligent harm is extended to researchers with an employment contract (substantive or honorary) with the Trust. The Trust only accepts liability for research activity that has been managerially approved by the R&D Department.

The Trust therefore accepts liability for the above research project and extends indemnity for negligent harm to cover you as principal investigator and the researchers listed on the Site Specific Information form. Should there be any changes to the research team please ensure that you inform the R&D Department and that s/he obtains an employment contract with the Trust if required.

Yours sincerely

Dr D R Norfolk
Associate Director of R&D

Approved documents
The documents reviewed and approved are listed as follows

Document	*Version*	*Date of document*
NHS R&D Form	2.0	
SSI Form	2.0	21/04/09
Protocol	1.0	06/02/09
REC Letter confirming favourable opinion		01/04/09
Evidence of Insurance		02/10/08
Sponsor Letter		06/02/09
Patient information sheet (LREC Approved)	4.0	23/03/09
Consent form (LREC Approved)	2.0	23/03/09

NHS
Bradford and Airedale

NHS Bradford and Airedale
Research management and governance support team
Clinical Quality
Level 2
Douglas Mill, Bowling Old Lane
Bradford BD5 7JR

OurRef://RMG/Approval/approval_letter_version_3

Monday, 8th June 2009

Ms Leila Mehdizadeh
23 Kingswear Garth
Whitkirk
Leeds
LS15 8LS

Re: Clinicians Views on Effective Clinical Decision making: An Interview Study

Ref no: 001_19_06_09_0000

Thank you for your recent submission to NHS Bradford and Airedale research management and governance support team.

Following consideration of your submission I am pleased to confirm that research management and governance approval has been granted by NHS Bradford and Airedale for the above research to take place as described in your completed application and accompanying documentation.

Conditions of approval

You should be aware that approval is granted subject to the conditions specified below:

- Throughout the course of the study, all research activity should comply with relevant, current governance and regulatory requirements including (but not limited to)
    - The Research Governance Framework for Health and Social Care, 2nd Ed (2005)
    - The Medicines for Human Use (Clinical Trials) Regulations (2004) and subsequent amendments
    - The Mental Capacity Act (2005)
    - The Ionising Radiation (Medical Exposure) (Amendment) Regulations (2006)
    - The Medical Devices Regulations (2002) (Statutory Instrument 2002/618)
    - The Human Tissue Act (2004)
    - The Data Protection Act (1998)

Q:\Clinical Quality & Governance\Clinical Quality\Research\Research_Management_Governance_May_2009\Approvals\09_10\Approvals_Q1_Apr_Jul_09\001_19_06_09_0000_Mehdizadeh.doc

NHS
Bradford and Airedale

- Consent for NHS Bradford and Airedale to audit your project, which is implicit in your acceptance of approval.

- Where any amendments, substantial or non substantial are made throughout the course of the study these should be notified to NHS Bradford and Airedale on the relevant form (available from http://myresearchproject.org)

- A copy of the final study report should be forwarded to NHS Bradford and Airedale on the relevant form (available from http://myresearchproject.org) no later than 3 months following study completion

- Should any serious adverse event(s) occur throughout the course of the study these should be notified to NHS Bradford and Airedale using the contact details set out above

Should you require any clarification regarding any of the points raised above, or have any further queries in relation to approvals and post approval study management process then please do not hesitate to contact me on 01274 237397.

Finally, may I take this opportunity to wish you well with your study and look forward to hearing about your progress in due course.

Yours sincerely

Dr Peter Dickson
Medical Director

NHS Bradford and Airedale

Ms Claire Seymour
Head of Quality Development

NHS Bradford and Airedale

Encs.

CC:

Ms Anna Frearson NHS Leeds

Q:\Clinical Quality & Governance\Clinical Quality\Research\Research_Management_Governance_May_2009\Approvals\09_10\Approvals_Q1_Apr_Jul_09\001_19_06_09_0000_Mehdizadeh.doc

8.7. Study information sheet

DOCTORS' VIEWS ABOUT MAKING CLINICAL DECISIONS EFFECTIVELY: AN INTERVIEW STUDY

Introduction

Making clinical decisions well is a key aspect of doctor's everyday practice. Little is known about how doctors learn to make clinical decisions well. This study asks doctors for their views and perceptions about what enhances clinical decision making. The study is being carried out by Leila Mehdizadeh, a doctoral student at the University of Leeds supervised jointly between the Leeds Institutes of Health Sciences (LIHS) and Medical Education (LIME).

Why am I being contacted?

Doctors working at the Leeds Teaching Hospitals or Leeds Primary Care Trust are being invited to take part in the study. This study is an interview study, requiring a sample of about 20 doctors. We would like this sample to include doctors at different stages in their careers from different specialties to make sure a range of views and experiences about clinical decision making are represented in the interviews.

What does the study involve?

If you agree to take part, Leila will arrange to meet with you to talk about your views and experiences of making decisions. Leila is happy to meet with you at your workplace; the interview lasts about 30 minutes. The interviews will be digitally recorded to ensure your views are accurately represented. There are no right and wrong answers to the questions Leila will ask you. It is your views and experiences that are important.

What happens to the interview data?

A third party will transcribe the digital recordings, making sure the resulting transcript is anonymised, i.e. your name will be replaced by a study identification number. Leila will use the anonymised paper and electronic versions of these transcripts when analysing their content. The content of your interview will be analysed together with those from other study participants. Anonymised quotes from these combined analyses will be used in the study results and disseminated via her thesis, publication and conference presentations.

Confidentiality and anonymity

Your identity will be anonymised so that no one can recognise you from the interview. The data from the interviews are confidential. The study materials will be stored in a locked filing cabinet in Leila's office in LIHS. Any electronic versions of the data will be erased when Leila completes her doctoral studies. The person transcribing the recordings will erase any versions of the recordings and transcripts on their computers after the interview is transcribed.

Is participation voluntary?

Yes, it is entirely your choice whether you participate or not. If you participate in the study, you can choose not to talk about an issue raised during the interview or you can withdraw from the study at any time without giving a reason.

What happens next?

If you would like more details about this study and/or are interested in participating then please contact Leila Mehdizadeh either via her email (psc3lm@leeds.ac.uk), mobile phone **(07765 911305)** or address (Leeds Institute of Health Sciences; Charles Thackrah Building - room G.02; University of Leeds; 101 Clarendon Road; Leeds LS2 9LJ).

Thank you for taking the time to read this information.

Dr Hilary Bekker ***Senior Lecturer in Behavioural Sciences; LIHS***

Dr Vikram Jha ***Consultant Obstetrician and Senior Lecturer in Medical Education; LIME***

Dr Naomi Quinton ***Lecturer in medical Education; LIME***

8.8. Consent form

Title of Project: Doctors' views about making clinical decisions effectively: an interview study

Name of Researcher: Leila Mehdizadeh

Please initial box

1. I confirm that I have read and understand the information sheet dated 23/03/09 (version 4.0) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and I can terminate the interview and/or withdraw from the study at any time without giving a reason, without my work or legal rights being affected. ☐

3. I permit the researcher to audio-record the interview and use anonymised extracts of data when reporting the study. ☐

4. I agree to take part in the above study. ☐

Name of participant ______________________________

Date ______________________________

Signature ______________________________

Name of person taking consent ______________________________

Date ______________________________

Signature ______________________________

8.9. Interview schedule

Preliminary

- On a typical day what problems do your patients present with? In which situations do you feel you are making decisions about their care?

Main body; Examples from clinical practice

- What do you find challenging when managing patients?
- Can you think of an example of a patient who had a bad outcome but you felt comfortable with your clinical judgment.
- Can you think of an example of a patient who had a good outcome but you felt uncomfortable with an aspect of your clinical judgment.

What did you learn from that?

- Are there any examples of when you would have made a decision differently?
- Are there examples of valuable learning experiences from colleagues that helped improve your clinical judgement?

Examples from formal teaching

- Are there examples that stand out from your formal teaching that helped improve your clinical judgement (lecture, article, book)
- (How do you think doctors learn to be competent decision makers)?
- Is there anything at the place of work you think affects your decision making?
- (situation, experience, personal values, something about the patient, working in a team vs alone, NHS budget)

8.10. Coding frame of themes and categories

Theme 1: Factors influencing judgements	Theme 2: Factors influencing decisions	Theme 3: Making decisions	Theme 4: Novice to expert	Theme 5: Professional practice	Theme 6: Developing decision making skills
1.1 Patient's appearance/behaviour	2.1 Patient non-clinical factors (d-p relationship,personality, behaviour,preferences, events, beliefs,emotions, relatives)	3.1 Intuition	4.1 Doctor is expert	5.1 Relationships at work (doctor-patient, SDM, colleagues priorities,disagreement,detachment)	6.1 Clinical experience (practise skills on manikins and real patients,on job).
1.2 Information (amount, source of information/opinions of others)	2.2 Patient clinical factors (condition type,lifestyle,history,co-morb)	3.2 Problem solving (systematic thinking, filtering information)	4.2 Ideal vs reality	5.2 Quality of care (conscious effort,satisfaction,equal standard,compassion)	6.2 Colleagues (observing,role, models,anecdotes,discussion,feedback,supervision)
1.3 Condition	2.3 Doctor factors (prefs,limitations,personality,emotion,experience)	3.3 Pattern recognition	4.3 Seeking help	5.3 Bad outcomes (inevitable, mistakes, rectifying mistake)	6.3 Acquiring theoretical knowledge (formal teaching,applied assessments,scenarios,online,personal study)
1.4 Consequences (risk,importance, urgency, ethical implications,peer pressure)	2.4 Workplace factors (time,day,colleagues,logistics,peer pressure)	3.4 Habit	4.4 Experience of being doctor (supervision,difficulty, thinking style, change in culture over time, peer pressure,uncertain)	5.4 Safe practice (safety netting, cautious practices e.g. asking colleagues, clear instructions to patients, multiple treatments)	6.4 Professional norms (accepted procedures, ABC,guidelines)
1.5 Other	2.5 Other (complexity, multiple factors)	3.5 Information gathering (patient records, memory, questioning, observations, examination, investigations).	4.5 Transition phase (reflections of seniors, difficulty for novices)	5.5 Reflective practice (awareness of decision process, good/bad outcomes, doctors' limitations, justifying reasoning)	6.5 Mistakes (or emotional experiences)
		3.6 Strategies (prioritising,forward planning,negotiating,guidelines,colleagues)	4.6 Other	5.6 Other	6.6 Other
		3.7 Other			

8.11. Demographic information of interview sample

Participant number	Gender	Current clinical specialty	Seniority level	Academic role	Educator role
2	Male	General Practice	Experienced 23 years	No	Yes
3	Male	General Practice	Experienced 18 years	Yes	Yes
4	Female	General practice	Middle grade 4 years	Yes	Yes
5	Male	Oncology	FY2	Yes	Yes
6	Male	Paediatric surgery	Junior registrar (ST1)	No	No
7	Female	Obstetrics and Gynaecology	Senior registrar	No	Yes
8	Male	Paediatric medicine	Consultant	No	Yes
9	Male	Renal medicine	Consultant	Yes	Yes
10	Female	Anaesthetics	Consultant	Yes	Yes
11	Male	Adult Psychiatry	Consultant	Yes	Yes
12	Female	N/A	5th year Undergraduate student	No	No
13	Female	N/A	5th year Undergraduate student	No	No
14	Male	N/A	4th year Undergraduate student	No	No
15	Female	Acute general medicine	FY1	No	No
16	Female	Vascular surgery	Junior registrar (ST1)	No	No

8.12. EdREC ethical approval for quasi-experiment

Faculty of Medicine and Health
Research Office

Room 10.110, Level 10
Worsley Building
Clarendon Way
Leeds LS2 9NL

T (General Enquiries) +44 (0) 113 343 4361
F +44 (0) 113 343 4373

UNIVERSITY OF LEEDS

Ms Leila Mehdizadeh
Leeds Institute of health Sciences
Room G.02, Charles Thackrah Building
University of Leeds
101 Clarendon Road
LEEDS LS2 9LJ

29th June 2010

Dear Leila

Re: EDREC/09/016
Title: Enhancing Medical Students' Awareness of Decision Making: A Pilot Experimental Study

I am pleased to inform you that the amendment for the above research application has been reviewed by the EdREC committee and following receipt of evidence of permissions requested, I can confirm a favourable ethical opinion.

Please notify the committee if you intend to make any further amendments to the original research as submitted at date of this approval prior to implementation.

I wish you every success with the project.

Yours sincerely

Professor Deborah Murdoch-Eaton
Chair, EdREC

8.13. Study information sheet

INFORMATION SHEET

Study title: *Enhancing doctors' awareness of decision making: a pilot experimental study*
Researchers: *Leila Mehdizadeh, Hilary Bekker, Vikram Jha, Naomi Quinton*

1. What is the purpose of this study?

Making clinical decisions well is a key aspect of doctor's everyday practice. Despite this, there is limited research evidence on how to help doctors acquire good clinical decision making skills. There is also little formal decision making training in the current undergraduate curriculum. The purpose of this study is to determine whether a short online tutorial is enough to 1) enhance doctors' awareness of people's decision making and 2) improve their decision making skills.

This experimental online study is being carried out by Leila Mehdizadeh, a PhD student at the University of Leeds supervised jointly between the Leeds Institutes of Health Sciences (LIHS) and Medical Education (LIME).

Students who are taking postgraduate courses within the School of Medicine and are also qualified medical doctors are being invited to take part in the study.

2. Do I have to take part?

It is entirely your choice whether you participate or not. If you participate, you can choose to withdraw from the study at any time without giving a reason. If you do not participate, then this has no impact on your studies.

3. How do I complete the study?

The study is completed online, as long as you have access to a computer with internet then you can participate from a location of your choice. You will be asked to complete a series of short tasks at your own pace, but the study is not expected to take longer than 15 minutes in total. Instructions will appear on screen to guide you throughout the study. The study includes 1) a decision making questionnaire based on four clinical scenarios 2) a short tutorial on how people make decisions and 3) an evaluation questionnaire.

4. Will my taking part in this study be kept confidential?

Yes, the data you provide will remain confidential and strictly for research purposes only. Participants' data will be stored in anonymous files on password protected computers that only Leila has access to. Your identity will remain anonymous at all times.

5. What's in it for me?

All participants will be entered in to a prize draw with the chance to win a leisure voucher. Leisure vouchers can be used in a large range of shops, restaurants, gyms and attractions (e.g. HMV, Waterstones, Pizza Hut, Alton Towers, Madame Tussauds, London Eye). After the study's closing date has passed, three people will be selected at random to each win a voucher. First prize will be a voucher of £20, second prize £10 and third prize £5.

As a doctor you will have to make many clinical decisions so you need to be confident that you make decisions in the best way possible. This study should help you develop an awareness of i) the importance of decision making as a key clinical skill and ii) how doctors make decisions in practice. Other anticipated benefits are that you reflect more on your own decision making and learn to make better decisions in practice. You might also find the tasks and tutorial interesting.

6. Results of the study?

The data will be discussed with Leila's supervisors but at no point will your identity be divulged. The study results will be disseminated via her PhD thesis, conference presentations and publication.

7. Who has reviewed the study?

The study has been reviewed by the University's Medicine and Dentistry Educational Research Ethics Committee (EdREC).

8. Consent

Participants are required to complete the 'informed consent' section shown after this information in order to proceed with the study. This is to confirm that you understand what the study involves and that you can withdraw at any point.

9. Contact details for further information:

If you would like more details about this study then please contact Leila Mehdizadeh either via her email (psc3lm@leeds.ac.uk), or address (Leeds Institute of Health Sciences; Charles Thackrah Building - room G.02; University of Leeds; 101 Clarendon Road; Leeds LS2 9LJ).

10. Closing date of study

The study is open for participation until ***30th June 2010.*** After this date the study will close and thereafter the winners of the prize draw will be contacted

INFORMED CONSENT (completed online)

Before you can begin the study, you are required to complete the following section that confirms you agree to take part. You cannot proceed without completing **all** of the information below.

Title of Project: Enhancing doctors' awareness of decision making: a pilot experimental study

Principle Researcher: Leila Mehdizadeh

Please select

1. I confirm that I have read and understand the information sheet. ☐

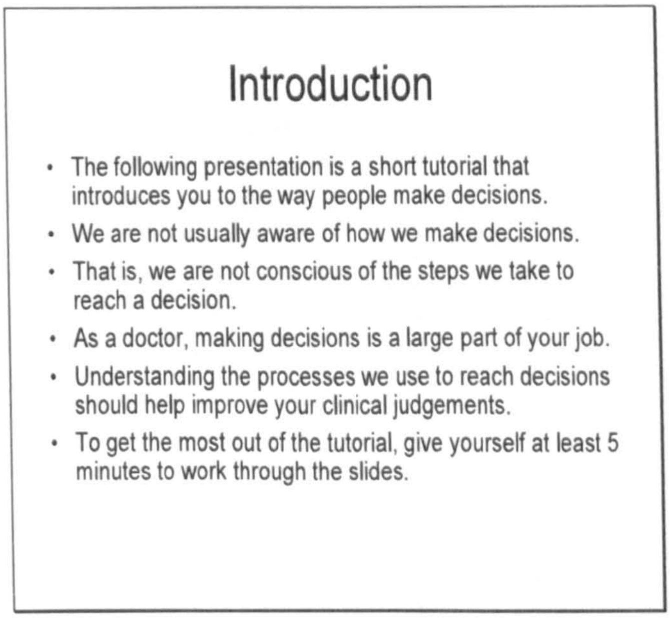
2. I confirm that I have had the opportunity to ask questions. ☐

3. I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason. ☐

4. I agree to take part in the above study. ☐

8.14. Tutorial PowerPoint slides

Slide 1



Slide 2

Slide 3

Aim & Learning Objectives

Aim: To enhance your awareness of how people make decisions.

Learning objectives:

- To understand that people generally make decisions in the same way as each other;
- To understand how decisions are made using a heuristic and/or systematic strategy;
- To understand the advantages and disadvantages of the heuristic and systematic strategy when making decisions;
- To be able to identify whether decisions have been made using a heuristic strategy or systematic strategy.

Slide 4

- Everyone tends to make decisions in the same way.
- Just as the human heart, kidneys, or lungs function in the same way from person to person, so does the brain.
- The brain has a given infrastructure that makes sense of information "out there".
- Actively making sense of information is known as **Information Processing.**
- Most of the time we are not aware of how we process information because it happens sub-consciously.
- We make our decisions based on the information we have processed and not on the full information that is available "out there".

Slide 5

Selecting Information

- We are always solving problems and making decisions.
- But there is so much complex information "out there" that we do not process all of it when making a decision.
- The brain has a **limited capacity** to process information consciously.
- The brain has ways of reducing what information we attend to, either based on our own experiences or external cues.

Limited capacity

Selecting information

Slide 6

Two Methods of Decision Making

❶ heuristic ❷ systematic

We all use two strategies to reach a decision;

❶ Heuristic strategy: Decisions are based on a subset of information using mental shortcuts.

❷ Systematic strategy: Decisions are based on all the known options and weighing up the advantages and disadvantages of their consequences.

Slide 7

Heuristic Strategy:

Taking Shortcuts

- Heuristics are simple rules of thumb that help us take shortcuts when we want to make judgements or decisions.
- For example, if you were asked if a baby was a boy or girl and it was dressed in pink, you would be likely to say "girl".
- Or when choosing a new camera you might buy the one recommended to you by a trusted friend rather than spend time searching for the latest model or best value for money.
- We make most of our decisions using heuristic methods.
- Heuristic strategies are quick and enables us to make decisions without much (if any) conscious effort.

The image below illustrates an example of a heuristic in use. Miss Tonks represents the shortcut, who will provide a quicker solution than the computer.

Slide 8

Heuristics and Clinical Decisions

- Doctors and patients use a type of expert heuristic to understand information and make decisions.
- You will experience lots of clinical scenarios that enable you to make quick and correct decisions.
- For example, you will be taught to always consider pre-eclampsia when a pregnant woman feels breathless and other 'worst case scenarios'.
- Or, you will follow the advice of a more experienced colleague if you are unsure of which type of treatment to give a patient.
- These are known as **smart heuristics** because they usually lead to correct decisions in an economical way.

Slide 9

Disadvantage of Shortcuts

- •Heuristics will often help you make good clinical decisions.
- •But using heuristics means you don't analyse details of information, so you are more likely to make a mistake.
- •Our judgements can be biased by the external and internal cues we use to focus our attention.
- •Examples of external cues of a situation are time pressure, time of day or the person that gives information.
- •A doctor is more willing to act on a colleague's opinion if there is no time to find out the correct answer for himself/herself. Or, a treatment decision can be more influenced by the free gifts from a sales representative instead of an evaluation of the evidence.
- •Internal cues are factors such as our beliefs, prejudices, experiences and habits.
- •For example, diagnostic decisions can be influenced by a doctor's beliefs about a patient e.g. "she's a hypochondriac there is nothing wrong with her" or past experiences tell us that "young and fit people don't have heart attacks". Doctors also develop habits such as a set of preferred drugs they always prescribe to patients.
- •These external and internal cues influence the shortcuts we use to make judgements.
- •The disadvantage is that we use shortcuts without questioning their appropriateness which can lead to incorrect judgements and decisions.

*The image on the right shows the doctor's treatment decision is influenced by the drug commercials he likes. This is **not** the best way to make a clinical decision!*

Slide 10

Systematic Strategy: The Longer Way Round

- Systematic strategies require conscious attention to the details of the options.
- It is a thorough evaluation of the advantages and disadvantages of the consequences of each option.
- Systematic processing is sometimes used to reach decisions that have important consequences for us.
- It is unlikely we would use a systematic strategy for what to eat for dinner tonight, you would more likely use a heuristic such as "whatever is there".
- But, when choosing a university course we might consider the reputation of the university, course quality, location, living expenses, recreational activities etc and choose the one that best meets our preferences and priorities.
- A systematic method of making decisions feels challenging to us as it requires conscious thought and is time consuming.

Slide 11

Thinking Systematically About Clinical Decisions

- Doctors will think systematically about most novel clinical decisions.
- For example, a doctor may weigh up the risks and benefits of performing or not performing high risk surgery on a particular patient.
- Or if a patient is not responding to a treatment in the way they should, a doctor may revisit the details and think of alternative diagnoses.
- The disadvantage of this method is the conscious effort required to think carefully about the information before deciding what to do.
- The advantage of systematic processing is there is less chance of making a wrong decision.
- Decisions made systematically are more stable i.e. they are less likely to change or be influenced by others.
- Also, we usually do not regret decisions we have made systematically because we can justify how we reached them.

Slide 12

Comparing Heuristic and Systematic Method of Decision Making

Heuristic	Systematic
Conscious effort is minimal	Conscious effort is required
No attention to details of choice	Attention to details of all options
Quick process	Time consuming
Feels easy	Feels challenging
Errors are likely	Errors are unlikely
Decisions not easily justified to others	Decisions easily justified to others
Unstable judgements (easily swayed by others opinions)	More stable judgements (not easily swayed by others opinions)

Slide 13

Take Home Message!

- You will use heuristic strategies when making clinical decisions. These mental shortcuts are essential and will lead to satisfactory decisions most of the time.

- Nevertheless they are shortcuts so be aware that they can sometimes lead to unsatisfactory and/or incorrect clinical decisions.

- Sometimes it is better to take a systematic approach and think things through methodically. It requires more effort to reach a decision this way, but it is more reliable, you are less likely to regret your decision and you are more likely to justify it to others when it is made systematically.

Slide 14

Summary

- Actively making sense of information is called information processing.
- People process information selectively because the brain has a limited capacity for conscious attention.
- There are two methods of information processing that people use to make decisions.
- Heuristic processing strategy is the use of mental shortcuts to make decisions quickly and without much effort.
- Smart heuristics are shortcuts based on experience that lead us to satisfactory decisions in an economical way.
- Systematic processing strategy requires consideration of all the known options and weighing up the advantages and disadvantages of their consequences.

8.15. Knowledge measure using multiple choice questions

The purpose of this questionnaire is to evaluate your knowledge and understanding of the information presented in the tutorial 'How do we make Decisions?' There are three parts to the questionnaire, please complete all of them.

PART 1

Below are 10 multiple choice questions. Please answer each question by placing a tick next to one option only.

1. Which of the following terms refers to actively making sense of information?	select one option
Information processing	
Limited capacity	
Heuristic and systematic strategy	
Heuristic strategy	
Smart heuristic	
Systematic strategy	

2. Which of the following terms explains why we do not process all of the information "out there"	select one option
Information processing	
Limited capacity	
Heuristic and systematic strategy	
Heuristic strategy	
Smart heuristic	
Systematic strategy	

3. Which of the following terms refers to the method we use to make decisions based on a subset of information using mental shortcuts?	select one option
Information processing	
Limited capacity	
Heuristic and systematic strategy	
Heuristic strategy	
Smart heuristic	
Systematic strategy	

4. Which of the following terms refers to the mental shortcuts that lead us to correct decisions in an economical way?	select one option
Information processing	
Limited capacity	
Heuristic and systematic strategy	
Heuristic strategy	
Smart heuristics	
Systematic strategy	

5. Which of the following terms refers to the method we use to reach decisions based on all known options and weighing up the advantages and disadvantages of their consequences?	select one option
Information processing	
Limited capacity	
Heuristic and systematic strategy	
Heuristic strategy	
Smart heuristics	
Systematic strategy	

6. Which of the following can sometimes lead us to unsatisfactory and/or incorrect decisions?	select one option
Information processing	
Limited capacity	
Heuristic and systematic strategy	
Heuristic strategy	
Smart heuristics	
Systematic strategy	

7. Which of the following refers to the mechanism that leads us to stable judgements (i.e. not easily swayed by others)	select one option
Information processing	
Limited capacity	
Heuristic and systematic strategy	
Heuristic strategy	
Smart heuristics	
Systematic strategy	

8. Which of the following are you most likely to use when deciding which film to go see at the cinema?	select one option
Information processing	
Limited capacity	
Heuristic and systematic strategy	
Heuristic strategy	
Smart heuristics	
Systematic strategy	

9. Which of the following are you most likely to use when deciding to purchase a house?	select one option
Information processing	
Limited capacity	
Heuristic and systematic strategy	
Heuristic strategy	
Smart heuristics	
Systematic strategy	

10. Which of the following terms refer to how experts make their professional decisions?	select one option
Information processing	
Limited capacity	
Heuristic and systematic strategy	
Heuristic strategy	
Smart heuristics	
Systematic strategy	

8.16. Measure of clinical relevance

This final section is an evaluation of your thoughts and opinions towards the tutorial you saw on how we make decisions. Please read the following statements. For each statement place a tick in one box only.

Statements	Agree	Neither agree nor disagree	Disagree
I found the tutorial interesting			
It has helped me to think about my own decision making			
I have a basic understanding of how people make decisions			
I understand the relevance of the tutorial to me as a doctor			
I can identify examples of when people use heuristic and systematic strategies to make personal decisions			
I can identify examples of when doctors use heuristics and systematic strategies to make clinical decisions			
I understand how others can affect my judgements and choices			
I understand how my beliefs and experiences can affect my judgements and choices			
I feel more cautious about making clinical decisions			
I feel more confused about making clinical decisions			
I do not feel I have benefitted from the tutorial			
I think some formal teaching on how people make decisions should be included in medical training			

8.17. Clinical judgement measure using problem solving tasks

Below are four short medical scenarios. Read each one and answer the accompanying questions about your diagnosis. Please try and answer them all, do not miss any questions out.

SCENARIO 1

A 19 year old man comes to A&E late at night complaining he is struggling to breathe. He is a student and this has been his first week at university. You take his history and find out he had mild asthma as a child, but feels he has grown out of it with age. You examine him and he sounds wheezy and is taking quick short breaths. He states he has no chest pain and has had no recent illness.

1. If you had to make a diagnosis now, what would you choose? ***(select one option only)***

- Pneumonia
- Panic attack
- Asthma attack
- Heart attack

2. How certain are you that this is the correct diagnosis? ***(select one option only)***

- not at all certain (4)
- somewhat certain (3)
- fairly certain (2)
- absolutely certain (1)

3. Would you actively wish to exclude or diagnose another condition? ***(select one option only)***

- Pneumonia

- Panic attack
- Asthma attack
- Heart attack
- None of the above
- Other

4. Rate how much influence the following aspects of the scenario had on your chosen diagnosis. 1 = most influence, 4 = least influence.

- First week at university
- History of asthma
- Signs of wheezing and shortness of breath
- Absence of chest pain and illness

SCENARIO 2

A mother brings her 1 year old baby into the GP surgery. She explains her baby has developed a fever and bad cough yesterday which has worsened during the night. As baby cries it starts coughing and you notice it sounds like a bark. You take the baby's temperature and it is slightly raised. The breathing is noisy, but the baby does not appear to be struggling for breath. There is no rash on the body and throat and ears look normal.

1. Given these findings only, what diagnosis would you make? *(select one option only)*

- Meningitis
- Croup
- Flu
- Pneumonia

2. How certain are you that this is the correct diagnosis? *(select one option only)*

- not at all certain

- somewhat certain
- fairly certain
- absolutely certain

3. Would you actively wish to exclude or diagnose another condition? *(select one option only)*

- Meningitis
- Croup
- Flu
- Pneumonia
- None of the above
- Other

4. Rate how much influence the following aspects of the scenario had on your chosen diagnosis. 1 = most influence, 4 = least influence

- Age of patient
- Fever
- Barking cough
- Absence of other symptoms

SCENARIO 3

A young man is brought into accident and emergency (A&E) in the early hours of one morning. The police had found him slumped on the steps of a public library. He is unshaven, his clothes are dirty and is not fully conscious. He is unwilling to arouse himself. He seems confused and cannot respond with any clarity to the nurse's questions (Groopman, 2007 page 55).

1. If you had to make a diagnosis now, what would you choose? *(select one option only)*

- Alcohol excess
- Head Injury

- Epilepsy
- Diabetes

2. How certain are you that this is the cause? ***(select one option only)***

- not at all certain
- somewhat certain
- fairly certain
- absolutely certain

3. Would you actively wish to exclude or diagnose another condition? ***(select one option only)***

- Alcohol excess
- Head Injury
- Epilepsy
- Diabetes
- None of the above
- Other

4. Rate how much influence the following aspects of the scenario had on your chosen diagnosis. 1 = most influence, 4 = least influence

- Patient found slumped on library steps
- His unkempt appearance (unshaven, dirty clothes)
- Not fully conscious
- Confused state

SCENARIO 4

A 41 year old man comes into accident and emergency (A&E) complaining of severe chest pain. He explains he was hiking in the woods when a pain in his chest stopped him in his tracks. He works as a forest ranger so due to his active lifestyle is trim and very fit.

Over the past few days he has experienced growing discomfort in his chest, even when resting. As a forest ranger he is used to muscle aches but thinks this is different. He has no history of heart problems (Groopman, 2007 page 42-43).

1. If you had to make a diagnosis now, what would you choose? ***(select one option only)***

- Muscle strain
- Indigestion
- Angina
- Anxiety

2. How certain are you that this is the most likely diagnosis? ***(select one option only)***

- not at all certain
- somewhat certain
- fairly certain
- absolutely certain

3. Would you actively wish to exclude or diagnose another condition? ***(select one option only)***

- Muscle strain
- Indigestion
- Angina
- Anxiety
- None of the above
- Other

4. Which piece of information in the scenario influenced your decision the most? ***Rank order each option using the numbers 1-4 where 1 = most influence and 4 = least influence.***

- His age

- Patient is active and physically fit
- Chest pain even at rest
- No history of heart problems

8.18. Examples of cross-tabulations and statistical tests

8.18.1. Diagnostic choice before vs after

Diagnostic choice before (asthma scenario) * Diagnostic choice after (asthma scenario) Crosstabulation

			Diagnostic choice after (asthma scenario)		
			panic	asthma attack	Total
Diagnostic choice before (asthma scenario)	panic	Count	2	5	7
		% within Diagnostic choice before (asthma scenario)	28.6%	71.4%	100.0%
		% within Diagnostic choice after (asthma scenario)	40.0%	11.6%	14.6%
		% of Total	4.2%	10.4%	14.6%
	asthma attack	Count	3	37	40
		% within Diagnostic choice before (asthma scenario)	7.5%	92.5%	100.0%
		% within Diagnostic choice after (asthma scenario)	60.0%	86.0%	83.3%
		% of Total	6.3%	77.1%	83.3%
	heart attack	Count	0	1	1
		% within Diagnostic choice before (asthma scenario)	.0%	100.0%	100.0%
		% within Diagnostic choice after (asthma scenario)	.0%	2.3%	2.1%
		% of Total	.0%	2.1%	2.1%
Total		Count	5	43	48
		% within Diagnostic choice before (asthma scenario)	10.4%	89.6%	100.0%
		% within Diagnostic choice after (asthma scenario)	100.0%	100.0%	100.0%
		% of Total	10.4%	89.6%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
McNemar-Bowker Test	.	.	.[a]
N of Valid Cases	48		

a. Computed only for a PxP table, where P must be greater than 1.

Diagnostic choice before (croup scenario) * Diagnostic choice after (croup sceario) Crosstabulation

			Diagnostic choice after (croup scneario)		Total
			croup	flu	
Diagnostic choice before (croup scenario)	croup	Count	46	1	47
		% within Diagnostic choice before (croup scenario)	97.9%	2.1%	100.0%
		% within Diagnostic choice after (croup scneario)	100.0%	50.0%	97.9%
		% of Total	95.8%	2.1%	97.9%
	flu	Count	0	1	1
		% within Diagnostic choice before (croup scenario)	.0%	100.0%	100.0%
		% within Diagnostic choice after (croup scneario)	.0%	50.0%	2.1%
		% of Total	.0%	2.1%	2.1%
Total		Count	46	2	48
		% within Diagnostic choice before (croup scenario)	95.8%	4.2%	100.0%
		% within Diagnostic choice after (croup scneario)	100.0%	100.0%	100.0%
		% of Total	95.8%	4.2%	100.0%

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		1.000[a]
N of Valid Cases	48	

a. Binomial distribution used.

Diagnostic choice before (diabetes scenario) * Diagnostic choice after (diabetes scenario) Crosstabulation

			Diagnostic choice after (diabetes scenario)			
			alcohol excess	head injury	diabetes	Total
Diagnostic choice before (diabetes scenario)	alcohol excess	Count	35	2	1	38
		% within Diagnostic choice before (diabetes scenario)	92.1%	5.3%	2.6%	100.0%
		% within Diagnostic choice after (diabetes scenario)	97.2%	18.2%	100.0%	79.2%
		% of Total	72.9%	4.2%	2.1%	79.2%
	head injury	Count	1	9	0	10
		% within Diagnostic choice before (diabetes scenario)	10.0%	90.0%	.0%	100.0%
		% within Diagnostic choice after (diabetes scenario)	2.8%	81.8%	.0%	20.8%
		% of Total	2.1%	18.8%	.0%	20.8%
Total		Count	36	11	1	48
		% within Diagnostic choice before (diabetes scenario)	75.0%	22.9%	2.1%	100.0%
		% within Diagnostic choice after (diabetes scenario)	100.0%	100.0%	100.0%	100.0%
		% of Total	75.0%	22.9%	2.1%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
McNemar-Bowker Test	.	.	.[a]
N of Valid Cases	48		

a. Computed only for a PxP table, where P must be greater than 1.

Diagnostic choice before (angina scenario) * Diagnostic choice after (angina scenario) Crosstabulation

			Diagnostic choice after (angina scenario)				
			muscle strain	indgestion	angina	anxiety	Total
Diagnostic choice before (angina scenario)	muscle strain	Count	5	0	2	0	7
		% within Diagnostic choice before (angina scenario)	71.4%	.0%	28.6%	.0%	100.0%
		% within Diagnostic choice after (angina scenario)	83.3%	.0%	5.6%	.0%	14.6%
		% of Total	10.4%	.0%	4.2%	.0%	14.6%
	indgestion	Count	0	5	4	0	9
		% within Diagnostic choice before (angina scenario)	.0%	55.6%	44.4%	.0%	100.0%
		% within Diagnostic choice after (angina scenario)	.0%	100.0%	11.1%	.0%	18.8%
		% of Total	.0%	10.4%	8.3%	.0%	18.8%
	angina	Count	1	0	30	0	31
		% within Diagnostic choice before (angina scenario)	3.2%	.0%	96.8%	.0%	100.0%
		% within Diagnostic choice after (angina scenario)	16.7%	.0%	83.3%	.0%	64.6%
		% of Total	2.1%	.0%	62.5%	.0%	64.6%
	anxiety	Count	0	0	0	1	1
		% within Diagnostic choice before (angina scenario)	.0%	.0%	.0%	100.0%	100.0%
		% within Diagnostic choice after (angina scenario)	.0%	.0%	.0%	100.0%	2.1%
		% of Total	.0%	.0%	.0%	2.1%	2.1%
Total		Count	6	5	36	1	48
		% within Diagnostic choice before (angina scenario)	12.5%	10.4%	75.0%	2.1%	100.0%
		% within Diagnostic choice after (angina scenario)	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	12.5%	10.4%	75.0%	2.1%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
McNemar-Bowker Test	4.333	2	.115
N of Valid Cases	48		

8.18.1.1. Asthma scenario

id they choose correct diagnosis before (asthma scenario) * Did they choose correct diagnos after (asthma scenario) Crosstabulation

			Did they choose correct diagnosis after (asthma scenario)		
			no	yes	Total
Did they choose correc diagnosis before (asthma scenario)	no	Count	8	0	8
		% within Did they choose correct diagnosis before (asthma scenario)	100.0%	.0%	100.0%
		% within Did they choose correct diagnosis after (asthma scenario)	66.7%	.0%	16.7%
		% of Total	16.7%	.0%	16.7%
	yes	Count	4	36	40
		% within Did they choose correct diagnosis before (asthma scenario)	10.0%	90.0%	100.0%
		% within Did they choose correct diagnosis after (asthma scenario)	33.3%	100.0%	83.3%
		% of Total	8.3%	75.0%	83.3%
Total		Count	12	36	48
		% within Did they choose correct diagnosis before (asthma scenario)	25.0%	75.0%	100.0%
		% within Did they choose correct diagnosis after (asthma scenario)	100.0%	100.0%	100.0%
		% of Total	25.0%	75.0%	100.0%

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		.125[a]
N of Valid Cases	48	

a. Binomial distribution used.

8.18.1.2. Croup scenario

'id they choose correct diagnosis before (croup scenario) * Did they choose correct diagnosis after (croup scenario) Crosstabulation

			Did they choose correct diagnosis after (croup scenario)		
			no	yes	Total
Did they choose correct diagnosis before (croup scenario)	no	Count	1	0	1
		% within Did they choose correct diagnosis before (croup scenario)	100.0%	.0%	100.0%
		% within Did they choose correct diagnosis after (croup scenario)	50.0%	.0%	2.1%
		% of Total	2.1%	.0%	2.1%
	yes	Count	1	46	47
		% within Did they choose correct diagnosis before (croup scenario)	2.1%	97.9%	100.0%
		% within Did they choose correct diagnosis after (croup scenario)	50.0%	100.0%	97.9%
		% of Total	2.1%	95.8%	97.9%
Total		Count	2	46	48
		% within Did they choose correct diagnosis before (croup scenario)	4.2%	95.8%	100.0%
		% within Did they choose correct diagnosis after (croup scenario)	100.0%	100.0%	100.0%
		% of Total	4.2%	95.8%	100.0%

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		1.000[a]
N of Valid Cases	48	

a. Binomial distribution used.

8.18.1.3. Diabetes scenario

Did they choose correct diagnosis before (diabetes scenario) * Did they choose correct diagnosis after (diabetes scenario) Crosstabulation

			Did they choose correct diagnosis after (diabetes scenario)		
			no	yes	Total
Did they choose correct diagnosis before (diabetes scenario)	no	Count	47	1	48
		% within Did they choose correct diagnosis before (diabetes scenario)	97.9%	2.1%	100.0%
		% within Did they choose correct diagnosis after (diabetes scenario)	100.0%	100.0%	100.0%
		% of Total	97.9%	2.1%	100.0%
Total		Count	47	1	48
		% within Did they choose correct diagnosis before (diabetes scenario)	97.9%	2.1%	100.0%
		% within Did they choose correct diagnosis after (diabetes scenario)	100.0%	100.0%	100.0%
		% of Total	97.9%	2.1%	100.0%

Warnings

No measures of association are computed for the crosstabulation of Did they choose correct diagnosis before (diabetes scenario) * Did they choose correct diagnosis after (diabetes scenario). At least one variable in each 2-way table upo which measures of association are computed is a constant.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
McNemar-Bowker Test	.	.	.[a]
N of Valid Cases	48		

a. Computed only for a PxP table, where P must be greater than 1.

8.18.1.4. Angina scenario

id they choose correct diagnosis before (angina scenario) * Did they choose correct diagnosi after (angina scenario) Crosstabulation

			Did they choose correct diagnosis after (angina scenario)		
			no	yes	Total
Did they choose correc diagnosis before (angina scenario)	no	Count	11	6	17
		% within Did they choose correct diagnosis before (angina scenario)	64.7%	35.3%	100.0%
		% within Did they choose correct diagnosis after (angina scenario)	91.7%	16.7%	35.4%
		% of Total	22.9%	12.5%	35.4%
	yes	Count	1	30	31
		% within Did they choose correct diagnosis before (angina scenario)	3.2%	96.8%	100.0%
		% within Did they choose correct diagnosis after (angina scenario)	8.3%	83.3%	64.6%
		% of Total	2.1%	62.5%	64.6%
Total		Count	12	36	48
		% within Did they choose correct diagnosis before (angina scenario)	25.0%	75.0%	100.0%
		% within Did they choose correct diagnosis after (angina scenario)	100.0%	100.0%	100.0%
		% of Total	25.0%	75.0%	100.0%

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		.125[a]
N of Valid Cases	48	

a. Binomial distribution used.

8.18.2. Certainty of diagnosis before vs after

8.18.2.1. Asthma scenario

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Certainty of diagnosis before (asthma scenario) * Certainty of diagnosis after (ashma scenario)	48	100.0%	0	.0%	48	100.0%

Certainty of diagnosis before (asthma scenario) * Certainty of diagnosis after (ashma scenario) Crosstabulation

			Certainty of diagnosis after (ashma scenario)				
			not at all certain	somewhat certain	fairly certain	absolutely certain	Total
Certainty of diagnosis before (asthma scenario)	not at all certain	Count	2	0	2	0	4
		% within Certainty of diagnosis before (asthma scenario)	50.0%	.0%	50.0%	.0%	100.0%
		% within Certainty of diagnosis after (ashma scenario)	50.0%	.0%	7.1%	.0%	8.3%
		% of Total	4.2%	.0%	4.2%	.0%	8.3%
	somewhat certain	Count	2	12	5	0	19
		% within Certainty of diagnosis before (asthma scenario)	10.5%	63.2%	26.3%	.0%	100.0%
		% within Certainty of diagnosis after (ashma scenario)	50.0%	92.3%	17.9%	.0%	39.6%
		% of Total	4.2%	25.0%	10.4%	.0%	39.6%
	fairly certain	Count	0	1	20	2	23
		% within Certainty of diagnosis before (asthma scenario)	.0%	4.3%	87.0%	8.7%	100.0%
		% within Certainty of diagnosis after (ashma scenario)	.0%	7.7%	71.4%	66.7%	47.9%
		% of Total	.0%	2.1%	41.7%	4.2%	47.9%
	absolutely certain	Count	0	0	1	1	2
		% within Certainty of diagnosis before (asthma scenario)	.0%	.0%	50.0%	50.0%	100.0%
		% within Certainty of diagnosis after (ashma scenario)	.0%	.0%	3.6%	33.3%	4.2%
		% of Total	.0%	.0%	2.1%	2.1%	4.2%
Total		Count	4	13	28	3	48
		% within Certainty of diagnosis before (asthma scenario)	8.3%	27.1%	58.3%	6.3%	100.0%
		% within Certainty of diagnosis after (ashma scenario)	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	8.3%	27.1%	58.3%	6.3%	100.0%

8.18.2.2. Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
Certainty of diagnosis after (ashma scenario) - Certainty of diagnosis before (asthma scenario)	Negative Ranks	4[a]	6.00	24.00
	Positive Ranks	9[b]	7.44	67.00
	Ties	35[c]		
	Total	48		

a. Certainty of diagnosis after (ashma scenario) < Certainty of diagnosis before (asthma scenario)

b. Certainty of diagnosis after (ashma scenario) > Certainty of diagnosis before (asthma scenario)

c. Certainty of diagnosis after (ashma scenario) = Certainty of diagnosis before (asthma scenario)

Test Statistics[b]

	Certainty of diagnosis after (ashma scenario) - Certainty of diagnosis before (asthma scenario)
Z	-1.615[a]
Asymp. Sig. (2-tailed)	.106

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

8.18.3. Croup scenario

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Certainty of diagnosis before (croup scenario) * Certainty of diagnosis after (croup scenario)	48	100.0%	0	.0%	48	100.0%

Certainty of diagnosis before (croup scenario) * Certainty of diagnosis after (croup scenario) Crosstabulation

			Certainty of diagnosis after (croup scenario)				
			not at all certain	somewhat certain	fairly certain	absolutely certain	Total
Certainty of diagnosis before (croup scenario)	not at all certain	Count	2	0	0	1	3
		% within Certainty of diagnosis before (croup scenario)	66.7%	.0%	.0%	33.3%	100.0%
		% within Certainty of diagnosis after (croup scenario)	50.0%	.0%	.0%	11.1%	6.3%
		% of Total	4.2%	.0%	.0%	2.1%	6.3%
	somewhat certain	Count	2	7	4	0	13
		% within Certainty of diagnosis before (croup scenario)	15.4%	53.8%	30.8%	.0%	100.0%
		% within Certainty of diagnosis after (croup scenario)	50.0%	70.0%	16.0%	.0%	27.1%
		% of Total	4.2%	14.6%	8.3%	.0%	27.1%
	fairly certain	Count	0	3	19	4	26
		% within Certainty of diagnosis before (croup scenario)	.0%	11.5%	73.1%	15.4%	100.0%
		% within Certainty of diagnosis after (croup scenario)	.0%	30.0%	76.0%	44.4%	54.2%
		% of Total	.0%	6.3%	39.6%	8.3%	54.2%
	absolutely certain	Count	0	0	2	4	6
		% within Certainty of diagnosis before (croup scenario)	.0%	.0%	33.3%	66.7%	100.0%
		% within Certainty of diagnosis after (croup scenario)	.0%	.0%	8.0%	44.4%	12.5%
		% of Total	.0%	.0%	4.2%	8.3%	12.5%
Total		Count	4	10	25	9	48
		% within Certainty of diagnosis before (croup scenario)	8.3%	20.8%	52.1%	18.8%	100.0%
		% within Certainty of diagnosis after (croup scenario)	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	8.3%	20.8%	52.1%	18.8%	100.0%

8.18.3.1. Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
Certainty of diagnosis after (croup scenario) - Certainty of diagnosis before (croup scenario)	Negative Ranks	7[a]	8.00	56.00
	Positive Ranks	9[b]	8.89	80.00
	Ties	32[c]		
	Total	48		

a. Certainty of diagnosis after (croup scenario) < Certainty of diagnosis before (croup scenario)

b. Certainty of diagnosis after (croup scenario) > Certainty of diagnosis before (croup scenario)

c. Certainty of diagnosis after (croup scenario) = Certainty of diagnosis before (croup scenario)

Test Statistics[b]

	Certainty of diagnosis after (croup scenario) - Certainty of diagnosis before (croup scenario)
Z	-.688[a]
Asymp. Sig. (2-tailed)	.491

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

8.18.4. Diabetes scenario

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Certainty of diagnosis before (diabete scenario) * Certainty of diagnosis after (diabetes scenario)	48	100.0%	0	.0%	48	100.0%

Certainty of diagnosis before (diabete scenario) * Certainty of diagnosis after (diabetes scenario) Crosstabulation

			Certainty of diagnosis after (diabetes scenario)			
			not at all certain	somewhat certain	fairly certain	Total
Certainty of diagnosis before (diabete scenario)	not at all certain	Count	26	4	0	30
		% within Certainty of diagnosis before (diabete scenario)	86.7%	13.3%	.0%	100.0%
		% within Certainty of diagnosis after (diabetes scenario)	92.9%	23.5%	.0%	62.5%
		% of Total	54.2%	8.3%	.0%	62.5%
	somewhat certain	Count	2	10	3	15
		% within Certainty of diagnosis before (diabete scenario)	13.3%	66.7%	20.0%	100.0%
		% within Certainty of diagnosis after (diabetes scenario)	7.1%	58.8%	100.0%	31.3%
		% of Total	4.2%	20.8%	6.3%	31.3%
	fairly certain	Count	0	3	0	3
		% within Certainty of diagnosis before (diabete scenario)	.0%	100.0%	.0%	100.0%
		% within Certainty of diagnosis after (diabetes scenario)	.0%	17.6%	.0%	6.3%
		% of Total	.0%	6.3%	.0%	6.3%
Total		Count	28	17	3	48
		% within Certainty of diagnosis before (diabete scenario)	58.3%	35.4%	6.3%	100.0%
		% within Certainty of diagnosis after (diabetes scenario)	100.0%	100.0%	100.0%	100.0%
		% of Total	58.3%	35.4%	6.3%	100.0%

8.18.4.1. Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
Certainty of diagnosis after (diabetes scenario) - Certainty of diagnosis before (diabete scenario)	Negative Ranks	5[a]	6.50	32.50
	Positive Ranks	7[b]	6.50	45.50
	Ties	36[c]		
	Total	48		

a. Certainty of diagnosis after (diabetes scenario) < Certainty of diagnosis before (diabete scenario)

b. Certainty of diagnosis after (diabetes scenario) > Certainty of diagnosis before (diabete scenario)

c. Certainty of diagnosis after (diabetes scenario) = Certainty of diagnosis before (diabete scenario)

Test Statistics[b]

	Certainty of diagnosis after (diabetes scenario) - Certainty of diagnosis before (diabete scenario)
Z	-.577[a]
Asymp. Sig. (2-tailed)	.564

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

8.18.4.2. Angina scenario

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Certainty of diagnosis before (angina scenario) * Certainty of diagnosis after (angina scenario)	48	100.0%	0	.0%	48	100.0%

Certainty of diagnosis before (angina scenario) * Certainty of diagnosis after (angina scenario) Crosstabulation

			Certainty of diagnosis after (angina scenario)			Total
			not at all certain	somewhat certain	fairly certain	
Certainty of diagnosis before (angina scenario)	not at all certain	Count	15	6	0	21
		% within Certainty of diagnosis before (angina scenario)	71.4%	28.6%	.0%	100.0%
		% within Certainty of diagnosis after (angina scenario)	93.8%	31.6%	.0%	43.8%
		% of Total	31.3%	12.5%	.0%	43.8%
	somewhat certain	Count	1	9	5	15
		% within Certainty of diagnosis before (angina scenario)	6.7%	60.0%	33.3%	100.0%
		% within Certainty of diagnosis after (angina scenario)	6.3%	47.4%	38.5%	31.3%
		% of Total	2.1%	18.8%	10.4%	31.3%
	fairly certain	Count	0	3	7	10
		% within Certainty of diagnosis before (angina scenario)	.0%	30.0%	70.0%	100.0%
		% within Certainty of diagnosis after (angina scenario)	.0%	15.8%	53.8%	20.8%
		% of Total	.0%	6.3%	14.6%	20.8%
	absolutely certain	Count	0	1	1	2
		% within Certainty of diagnosis before (angina scenario)	.0%	50.0%	50.0%	100.0%
		% within Certainty of diagnosis after (angina scenario)	.0%	5.3%	7.7%	4.2%
		% of Total	.0%	2.1%	2.1%	4.2%
Total		Count	16	19	13	48
		% within Certainty of diagnosis before (angina scenario)	33.3%	39.6%	27.1%	100.0%
		% within Certainty of diagnosis after (angina scenario)	100.0%	100.0%	100.0%	100.0%
		% of Total	33.3%	39.6%	27.1%	100.0%

8.18.4.3. Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
Certainty of diagnosis after (angina scenario) - Certainty of diagnosis before (angina scenario)	Negative Ranks	6[a]	9.92	59.50
	Positive Ranks	11[b]	8.50	93.50
	Ties	31[c]		
	Total	48		

a. Certainty of diagnosis after (angina scenario) < Certainty of diagnosis before (angina scenario)

b. Certainty of diagnosis after (angina scenario) > Certainty of diagnosis before (angina scenario)

c. Certainty of diagnosis after (angina scenario) = Certainty of diagnosis before (angina scenario)

Test Statistics[b]

	Certainty of diagnosis after (angina scenario) - Certainty of diagnosis before (angina scenario)
Z	-.894[a]
Asymp. Sig. (2-tailed)	.371

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

8.18.5. Excluding or diagnosing other conditions

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Would you actively wis to exclude or diagnose another BEFORE (asthma scenario) * Would you actively wis to exclude or diagnose another AFTER (asthm scenario)	48	100.0%	0	.0%	48	100.0%

Would you actively wish to exclude or diagnose another BEFORE (asthma scenario) * Would you actively wish to exclude or diagnose another AFTER (asthma scenario) Crosstabulation

			Would you actively wish to exclude or diagnose another AFTER (asthma scenario)						Total
			pneumonia	panic attack	asthma attack	none of above	other	7.00	
Would you actively wish to exclude or diagnose another BEFORE (asthma scenario)	pneumonia	Count	2	0	0	1	0	3	6
		% within Would you actively wish to exclude or diagnose another BEFORE (asthma scenario)	33.3%	.0%	.0%	16.7%	.0%	50.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (asthma scenario)	50.0%	.0%	.0%	20.0%	.0%	11.5%	12.5%
		% of Total	4.2%	.0%	.0%	2.1%	.0%	6.3%	12.5%
	panic attack	Count	0	4	1	0	1	2	8
		% within Would you actively wish to exclude or diagnose another BEFORE (asthma scenario)	.0%	50.0%	12.5%	.0%	12.5%	25.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (asthma scenario)	.0%	44.4%	50.0%	.0%	50.0%	7.7%	16.7%
		% of Total	.0%	8.3%	2.1%	.0%	2.1%	4.2%	16.7%
	asthma attack	Count	0	3	1	0	0	0	4
		% within Would you actively wish to exclude or diagnose another BEFORE (asthma scenario)	.0%	75.0%	25.0%	.0%	.0%	.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (asthma scenario)	.0%	33.3%	50.0%	.0%	.0%	.0%	8.3%
		% of Total	.0%	6.3%	2.1%	.0%	.0%	.0%	8.3%
	none of above	Count	0	0	0	3	0	0	3
		% within Would you actively wish to exclude or diagnose another BEFORE (asthma scenario)	.0%	.0%	.0%	100.0%	.0%	.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (asthma scenario)	.0%	.0%	.0%	60.0%	.0%	.0%	6.3%
		% of Total	.0%	.0%	.0%	6.3%	.0%	.0%	6.3%
	other	Count	0	1	0	0	0	3	4
		% within Would you actively wish to exclude or diagnose another BEFORE (asthma scenario)	.0%	25.0%	.0%	.0%	.0%	75.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (asthma scenario)	.0%	11.1%	.0%	.0%	.0%	11.5%	8.3%
		% of Total	.0%	2.1%	.0%	.0%	.0%	6.3%	8.3%
	multiple conditions	Count	2	1	0	1	1	18	23
		% within Would you actively wish to exclude or diagnose another BEFORE (asthma scenario)	8.7%	4.3%	.0%	4.3%	4.3%	78.3%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (asthma scenario)	50.0%	11.1%	.0%	20.0%	50.0%	69.2%	47.9%
		% of Total	4.2%	2.1%	.0%	2.1%	2.1%	37.5%	47.9%
Total		Count	4	9	2	5	2	26	48
		% within Would you actively wish to exclude or diagnose another BEFORE (asthma scenario)	8.3%	18.8%	4.2%	10.4%	4.2%	54.2%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (asthma scenario)	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	8.3%	18.8%	4.2%	10.4%	4.2%	54.2%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
McNemar-Bowker Test	.	.	.[a]
N of Valid Cases	48		

a. Both variables must have identical values of categories.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Would you actively wish to exclude or diagnose another BEFORE (croup scenario) * Would you actively wish to exclude or diagnose another AFTER (croup scenario)	48	100.0%	0	.0%	48	100.0%

Would you actively wish to exclude or diagnose another BEFORE (croup scenario) * Would you actively wish to exclude or diagnose another AFTER (croup scenario) Crosstabulation

			Would you actively wish to exclude or diagnose another AFTER (croup scenario)						
			meningitis	flu	pneumonia	none of above	other	multiple conditions	Total
Would you actively wish to exclude or diagnose another BEFORE (croup scenario)	meningitis	Count	2	0	0	0	0	2	4
		% within Would you actively wish to exclude or diagnose another BEFORE (croup scenario)	50.0%	.0%	.0%	.0%	.0%	50.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (croup scenario)	100.0%	.0%	.0%	.0%	.0%	7.1%	8.3%
		% of Total	4.2%	.0%	.0%	.0%	.0%	4.2%	8.3%
	flu	Count	0	3	0	0	0	1	4
		% within Would you actively wish to exclude or diagnose another BEFORE (croup scenario)	.0%	75.0%	.0%	.0%	.0%	25.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (croup scenario)	.0%	60.0%	.0%	.0%	.0%	3.6%	8.3%
		% of Total	.0%	6.3%	.0%	.0%	.0%	2.1%	8.3%
	pneumonia	Count	0	0	3	0	0	6	9
		% within Would you actively wish to exclude or diagnose another BEFORE (croup scenario)	.0%	.0%	33.3%	.0%	.0%	66.7%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (croup scenario)	.0%	.0%	42.9%	.0%	.0%	21.4%	18.8%
		% of Total	.0%	.0%	6.3%	.0%	.0%	12.5%	18.8%
	none of above	Count	0	1	2	5	0	2	10
		% within Would you actively wish to exclude or diagnose another BEFORE (croup scenario)	.0%	10.0%	20.0%	50.0%	.0%	20.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (croup scenario)	.0%	20.0%	28.6%	100.0%	.0%	7.1%	20.8%
		% of Total	.0%	2.1%	4.2%	10.4%	.0%	4.2%	20.8%
	other	Count	0	0	0	0	1	0	1
		% within Would you actively wish to exclude or diagnose another BEFORE (croup scenario)	.0%	.0%	.0%	.0%	100.0%	.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (croup scenario)	.0%	.0%	.0%	.0%	100.0%	.0%	2.1%
		% of Total	.0%	.0%	.0%	.0%	2.1%	.0%	2.1%
	multiple conditions	Count	0	1	2	0	0	17	20
		% within Would you actively wish to exclude or diagnose another BEFORE (croup scenario)	.0%	5.0%	10.0%	.0%	.0%	85.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (croup scenario)	.0%	20.0%	28.6%	.0%	.0%	60.7%	41.7%
		% of Total	.0%	2.1%	4.2%	.0%	.0%	35.4%	41.7%
Total		Count	2	5	7	5	1	28	48
		% within Would you actively wish to exclude or diagnose another BEFORE (croup scenario)	4.2%	10.4%	14.6%	10.4%	2.1%	58.3%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (croup scenario)	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	4.2%	10.4%	14.6%	10.4%	2.1%	58.3%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
McNemar-Bowker Test	9.000	6	.174
N of Valid Cases	48		

ish to exclude or diagnose another BEFORE (diabetes scenario) * Would you actively wish to exclude or diagnose anot scenario) Crosstabulation

			tively wish to exclude or diagnose another AFTER (diabete						
			alcohol	ead injury	iabetes	ne of abov	other	multiple onditions	Total
Would you acti wish to exclude diagnose anoth BEFORE (diabet scenario)	head injury	Count	1	2	0	0	0	0	3
		% within Would you actively wish to ex diagnose another BEFORE (diabetes scenario)	33.3%	66.7%	.0%	.0%	.0%	.0%	100.0%
		% within Would you actively wish to ex diagnose another A (diabetes scenario	100.0%	100.0%	.0%	.0%	.0%	.0%	6.3%
		% of Total	2.1%	4.2%	.0%	.0%	.0%	.0%	6.3%
	diabetes	Count	0	0	1	0	0	1	2
		% within Would you actively wish to ex diagnose another BEFORE (diabetes scenario)	.0%	.0%	50.0%	.0%	.0%	50.0%	100.0%
		% within Would you actively wish to ex diagnose another A (diabetes scenario	.0%	.0%	50.0%	.0%	.0%	2.4%	4.2%
		% of Total	.0%	.0%	2.1%	.0%	.0%	2.1%	4.2%
	none of above	Count	0	0	1	0	0	0	1
		% within Would you actively wish to ex diagnose another BEFORE (diabetes scenario)	.0%	.0%	100.0%	.0%	.0%	.0%	100.0%
		% within Would you actively wish to ex diagnose another A (diabetes scenario	.0%	.0%	50.0%	.0%	.0%	.0%	2.1%
		% of Total	.0%	.0%	2.1%	.0%	.0%	.0%	2.1%
	other	Count	0	0	0	0	1	1	2
		% within Would you actively wish to ex diagnose another BEFORE (diabetes scenario)	.0%	.0%	.0%	.0%	50.0%	50.0%	100.0%
		% within Would you actively wish to ex diagnose another A (diabetes scenario	.0%	.0%	.0%	.0%	100.0%	2.4%	4.2%
		% of Total	.0%	.0%	.0%	.0%	2.1%	2.1%	4.2%
	multiple condi	Count	0	0	0	1	0	39	40
		% within Would you actively wish to ex diagnose another BEFORE (diabetes scenario)	.0%	.0%	.0%	2.5%	.0%	97.5%	100.0%
		% within Would you actively wish to ex diagnose another A (diabetes scenario	.0%	.0%	.0%	100.0%	.0%	95.1%	83.3%
		% of Total	.0%	.0%	.0%	2.1%	.0%	81.3%	83.3%
Total		Count	1	2	2	1	1	41	48
		% within Would you actively wish to ex diagnose another BEFORE (diabetes scenario)	2.1%	4.2%	4.2%	2.1%	2.1%	85.4%	100.0%
		% within Would you actively wish to ex diagnose another A (diabetes scenario	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	2.1%	4.2%	4.2%	2.1%	2.1%	85.4%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
McNemar-Bowker Test	.	.	.[a]
N of Valid Cases	48		

a. Computed only for a PxP table, where P must be greater than 1.

Would you actively wish to exclude or diagnose another BEFORE (angina scenario) * Would you actively wish to exclude or diagnose another AFTER (angina scenario) Crosstabulation

			Would you actively wish to exclude or diagnose another AFTER (angina scenario)						
			muscle strain	indigestion	angina	anxiety	none	multiple conditions	Total
Would you actively wish to exclude or diagnose another BEFORE (angina scenario)	indigestion	Count	0	1	0	0	0	1	2
		% within Would you actively wish to exclude or diagnose another BEFORE (angina scenario)	.0%	50.0%	.0%	.0%	.0%	50.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (angina scenario)	.0%	50.0%	.0%	.0%	.0%	2.9%	4.2%
		% of Total	.0%	2.1%	.0%	.0%	.0%	2.1%	4.2%
	angina	Count	0	0	2	0	0	3	5
		% within Would you actively wish to exclude or diagnose another BEFORE (angina scenario)	.0%	.0%	40.0%	.0%	.0%	60.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (angina scenario)	.0%	.0%	66.7%	.0%	.0%	8.8%	10.4%
		% of Total	.0%	.0%	4.2%	.0%	.0%	6.3%	10.4%
	anxiety	Count	0	0	0	1	0	1	2
		% within Would you actively wish to exclude or diagnose another BEFORE (angina scenario)	.0%	.0%	.0%	50.0%	.0%	50.0%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (angina scenario)	.0%	.0%	.0%	50.0%	.0%	2.9%	4.2%
		% of Total	.0%	.0%	.0%	2.1%	.0%	2.1%	4.2%
	none	Count	0	0	0	0	5	1	6
		% within Would you actively wish to exclude or diagnose another BEFORE (angina scenario)	.0%	.0%	.0%	.0%	83.3%	16.7%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (angina scenario)	.0%	.0%	.0%	.0%	83.3%	2.9%	12.5%
		% of Total	.0%	.0%	.0%	.0%	10.4%	2.1%	12.5%
	other	Count	0	0	0	0	1	2	3
		% within Would you actively wish to exclude or diagnose another BEFORE (angina scenario)	.0%	.0%	.0%	.0%	33.3%	66.7%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (angina scenario)	.0%	.0%	.0%	.0%	16.7%	5.9%	6.3%
		% of Total	.0%	.0%	.0%	.0%	2.1%	4.2%	6.3%
	multiple conditions	Count	1	1	1	1	0	26	30
		% within Would you actively wish to exclude or diagnose another BEFORE (angina scenario)	3.3%	3.3%	3.3%	3.3%	.0%	86.7%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (angina scenario)	100.0%	50.0%	33.3%	50.0%	.0%	76.5%	62.5%
		% of Total	2.1%	2.1%	2.1%	2.1%	.0%	54.2%	62.5%
Total		Count	1	2	3	2	6	34	48
		% within Would you actively wish to exclude or diagnose another BEFORE (angina scenario)	2.1%	4.2%	6.3%	4.2%	12.5%	70.8%	100.0%
		% within Would you actively wish to exclude or diagnose another AFTER (angina scenario)	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	2.1%	4.2%	6.3%	4.2%	12.5%	70.8%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
McNemar-Bowker Test	.	.	.[a]
N of Valid Cases	48		

a. Both variables must have identical values of categories.