Econometric analysis of unpaid carers' health-related quality of life for use in economic evaluation for healthcare decision making

Becky Pennington

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

Doctor of Philosophy

The University of Sheffield

Faculty of Medicine, Dentistry and Health

School of Medicine and Population Health

January 2026

Acknowledgements

This thesis has taken rather a long time since the inception of the idea in 2019 and has developed during a period of pandemic and two maternity leaves. As such, I am indebted to a lot of people for their support during these strange and stressful times.

I am very grateful to my supervisors Monica Hernandez-Alava and Mark Strong for their helpful discussions throughout this thesis. I would also like to thank Allan Wailoo for letting me run with a project on carer utilities back in 2018 which ultimately led to this research.

Undertaking a PhD has been somewhat lonely at times, and I would have gone slightly mad were it not for my officemates and colleagues, in particular Sarah Davis whose friendship and support over the past 8 years has been invaluable. Equally the meetings with NICE technology appraisal committee A have been a highlight of each month, interacting with other humans (although often via Zoom) and discussing different viewpoints.

I'd also like to thank colleagues and collaborators who have let me talk on and on about carer utilities even where we disagreed, including Hareth Al-Janabi, Paul Tappenden and Holly Cranmer.

I was initially terrified of speak to unpaid carers but this has proved to be illuminating, and so I am grateful to all of my public involvement contributors and facilitators, including Jill Carlton, Rosie Bamber, Catherine Cairns Leyman, Ella Fitzpatrick and Janine Wigmore.

I very much valued the experience of attending the Centre for Care Summer School in 2025 and all the mentors and other students I spoke to, whose research into different areas of care vastly broadened my understanding of care and care research.

The experience has been hugely brightened by my friend Gemma Shields concurrently undertaking a PhD so we could discuss the many common challenges and support each other. I am only slightly envious that she submitted before me.

I couldn't have finished this without support from my family, particularly my husband, parents and parents-in-law who believed I could do this but understood my expressions well enough to never ask too many questions about progress. And for looking after the children, especially where training courses and conferences have taken me away from home.

Lastly, my children, who have always been there for a cuddle and ridiculous ways to distract me, even if they have described my work on this PhD as "doing nothing all day."

Abstract

Economic evaluation is used to determine the most efficient use of resources by comparing the costs and outcomes of different interventions. Outcomes can be measured by health-related quality of life (HRQoL) and may include both patients and unpaid carers' HRQoL. Prior research found that incorporating carer's HRQoL in economic evaluations is uncommon, inconsistent, and relies on limited data and strong assumptions.

A key consideration is whether caregiving causally affects carer's HRQoL. My literature review indicates a lack of robust evidence. Generally, longitudinal studies suggest a detrimental effect on health/wellbeing, but the benefits derived from caregiving may mitigate its negative effects.

I analysed longitudinal data for patient-carer dyads and found carers' HRQoL decreased as the patients' HRQoL decreased and as the duration of care increased. My analysis of a cross-sectional EQ-5D dataset found people providing more hours of care per week have worse HRQoL, even after adjusting for confounders. Both analyses suggest a causal effect of caring on HRQoL and provide inputs for use in economic evaluations in any adult health condition.

Current methods for modelling carers' HRQoL assume the effect of caring on HRQoL is unidirectional: a negative effect implies any extension to patient survival leads to a carers' quality-adjusted life year (QALY) loss, a positive effect implies any extension to patient survival leads to a carers' QALY gain. These approaches force a choice between carers maximising their own HRQoL or the patients' HRQoL. My new modelling method allows a trade-off between the positive carer's HRQoL effect of improving patients' HRQoL and the negative carers' HRQoL effect of increasing caregiving burden.

I demonstrated that it is possible to identify a causal effect of caring and created a method to include this in economic evaluation. Future research can build on this to address data gaps and explore the transferability of evidence.

Table of Contents

List of Tables

List of Figures

Declaration

I, the author, confirm that the Thesis is my own work. I am aware of the University's Guidance on the Use of Unfair Means (*https://www.sheffield.ac.uk/study-skills/assessment/academic-integrity/academic-integrity*). This work has not previously been presented for an award at this, or any other, university.

Abbreviations

ADL	Activities of Daily Living
ASCOT-Carer	Adult Social Care Outcomes Toolkit-Carer
ATE	Average Treatment Effect
ATT	Average Treatment Effect on the Treated
CASP	Control, Autonomy, Self-realization, and Pleasure
CES-D	Center for Epidemiological Studies-Depression
CES	Carer Experience Scale
DDD	Difference-in-Difference-in-Differences
DID	Difference-in-Differences
DMD	Duchenne Muscular Dystrophy
EEPRU	Research Unit in Economic Methods of Evaluation of Health and Care Interventions
EQ-5D 3L	EQ-5D 3 level
EQ-5D 5L	EQ-5D 5 level
EQ-HWB	EQ Health and Wellbeing
GEE	General estimating equations
GMM	Generalised method of moments
HRQoL	Health-related quality of life
HST	Highly Specialised Technology
HTA	Health technology assessment
HUI	Health Utility Index
IADL	Instrumental activities of daily living
ICE-CAP A	ICEpop CAPability measure for Adults
ICER	Incremental cost-effectiveness ratio
IPW	Inverse probability weighting

MEPS	Medical Expenditure Panel Survey
NICE	National Institute for Health and Care Excellence
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary least squares
OR	Outcome regression
QALY	Quality-adjusted life year
SF-12	Short Form 12
SF-36	Short Form 36
SF-6D	Short Form 6 Dimension
SF MCS	Short Form Mental component score
SF PCS	Short Form Physical component score
SHEER	Spillovers in Health Economic Evaluation and Research
SMA	Spinal Muscular Atrophy
STREAM	Study on Transitions in Employment, Ability and Motivation
TA	Technology Appraisal
UKHLS	UK Household Longitudinal Survey

1 Introduction

Economic evaluation compares the costs and outcomes of an intervention to the costs and outcomes of one or more comparator(s) (alternative course(s) of action). In the case of healthcare interventions, the costs are usually those incurred by the health service, and outcomes are typically health outcomes for the patients receiving the intervention. Economic evaluation may be used to inform health technology assessment (HTA), in determining whether the healthcare intervention should be available for widespread use. In England, the National Institute for Health and Care Excellence (NICE) undertakes HTA, with economic evaluation as a key part of the process.

NICE specifies the methods to be used for HTA in its Guide to the Methods of Health Technology Evaluation (National Institute for Health and Care Excellence, 2022). This states that economic evaluation should take the form of cost-utility analysis, with outcomes expressed as health state utilities. Utilities value health-related quality of life (HRQoL) and are anchored between 0 and 1, where 0 is equivalent to death and 1 is full health (negative scores are possible, for health states considered worse than death). Health state utilities can be multiplied by the time spent in that health state to generate quality-adjusted life years (QALYs). The difference in costs between the intervention and comparator can be divided by the difference in QALYs to calculate the incremental cost-effectiveness ratio (ICER). Decision makers, such as NICE in England, can compare the ICER to a prespecified threshold to determine whether the intervention represents an efficient use of resources.

Healthcare interventions for patients can also affect health outcomes for the family members and friends who provide unpaid care to the patient (Al-Janabi et al, 2016a). There is therefore an argument for also including carers' health outcomes in economic evaluation of healthcare interventions for patients, since to exclude these would undermine the value of the contributions made by unpaid carers. As economist Nancy Folbre discussed in "The Invisible Heart" (2001), economic evaluation has traditionally assumed that altruism is a given, rather than attempting to value and include it:

> *"Economic theory offers us some indispensable tools for organizing our joint endeavours. That's why it's so important to get economics right – to ensure that it includes careful consideration of love, obligation and reciprocity, as well as self-interest. We must stop assuming that norms and preferences of caring for others come from "outside" our economic system and can therefore be taken as a given."* p120 (Folbre, 2001)

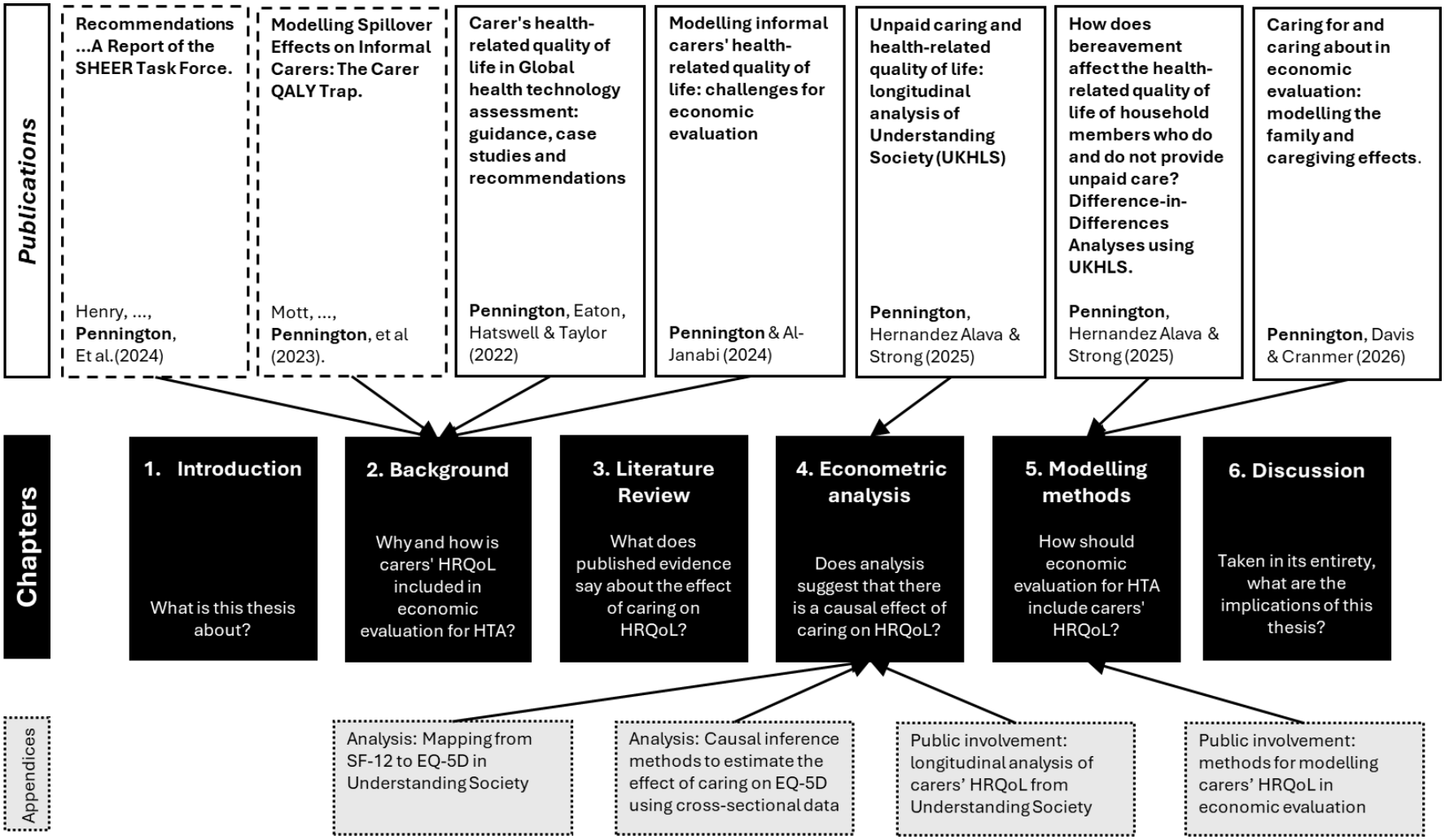
While the inclusion of carers' QALYs in economic evaluation is still relatively rare, it can have important implications for ICERs and associated reimbursement decisions. The extent to which carers' QALYs are affected by modelled patient interventions depends upon the intervention itself, the size of the carer utilities, and the method used to include carers' QALYs. In a review of NICE appraisals that included carers' HRQoL (undertaken prior to this PhD), I found that including carers' QALYs could change ICERs by up to 33% (Pennington, 2020). Given the potential for carers' QALYs to influence reimbursement decisions, it is vital to understand how to do so in a consistent and robust manner.

This work aimed to understand how carers' HRQoL changes over time, and to advance methods to include carers' HRQoL in economic evaluation. The objectives were:

1. To understand how economic evaluation currently includes carers' HRQoL
2. To review the current evidence to inform the effect of caring on HRQoL
3. To analyse the effect of caring on HRQoL, with particular focus on longitudinal data
4. To develop new methods for including the effect of caring on HRQoL in economic evaluation

This work has culminated in this thesis, the development of five first-author publications (all published and open access), and collaboration on two further publications.

An overview of the structure of the thesis, and the role of publications and appendices in each chapter is provided in Figure 1.



Black-filled: thesis chapter, White-filled: published work, Grey-filled: unpublished work.
Solid outline: first-authored publication, Dashed outline: co-authored publication, Dotted outline: appendix.

Figure 1: Thesis structure

1.1 Publications included in the thesis

The following publications are included in the thesis and referred to using the text in bold. The full citations (including doi to access full text papers free of charge) are provided here. These are the publications shown in white boxes in Figure 1.

Pennington et al, 2022

Carers' Health-Related Quality of Life in Global Health Technology Assessment: Guidance, Case Studies and Recommendations. Pennington B, Eaton J, Hatswell AJ, Taylor H. Pharmacoeconomics. 2022 Sep;40(9):837-850. doi: 10.1007/s40273-022-01164-4.

Pennington & Al-Janabi, 2024

Modelling Informal Carers' Health-Related Quality of Life: Challenges for Economic Evaluation. Pennington B & Al-Janabi H. Appl Health Econ Health Policy. 2024 Jan;22(1):9-16. doi: 10.1007/s40258-023-00834-4.

Pennington et al, 2025b

Unpaid Caring and Health-Related Quality of Life: Longitudinal Analysis of Understanding Society (the UK Household Longitudinal Survey). Pennington B, Hernández Alava M, Strong M. Value Health 2025 Jan;28(1):138-147. doi: 10.1016/j.jval.2024.08.004.

Pennington et al, 2025a

How Does Bereavement Affect the Health-Related Quality of Life of Household Members Who Do and Do Not Provide Unpaid Care? Difference-in-Differences Analyses Using the UK Household Longitudinal Survey. Pennington B, Hernández Alava M, Strong M. Pharmacoeconomics. 2025 Apr;43(4):389-402. doi: 10.1007/s40273-024-01452-1.

Pennington et al, 2026

Caring for and Caring about in Economic Evaluation: Modelling the Family and Caregiving Effects. Pennington B, Davis S, Cranmer H. Pharmacoeconomics. 2026 Jan;44(1):13-23. doi: 10.1007/s40273-025-01540-w

Mott et al (2023)

Modelling Spillover Effects on Informal Carers: The Carer QALY Trap. Mott D, Schirrmacher H, Al-Janabi H, Guest S, **Pennington B**, Scheuer N, Shah K, Skedgel C. Pharmacoeconomics. 2023 Dec;41(12):1557-1561. doi: 10.1007/s40273-023-01316-0.

Henry et al (2024)

Recommendations for Emerging Good Practice and Future Research in Relation to Family and Caregiver Health Spillovers in Health Economic Evaluations: A Report of the SHEER Task Force. Henry E, Al-Janabi H, Brouwer W, Cullinan E, Engel L, Griffin S, Hulme C, Kingkaew P, Lloyd A, Payakachat N, **Pennington B**, Peña-Longobardo L, Prosser L, Shah K, Ungar W, Wilkinson T, Wittenberg E. Pharmacoeconomics. 2024 Mar;42(3):343-362. doi: 10.1007/s40273-023-01321-3.

1.2 Chapters of the thesis

The thesis contains six chapters (plus appendices). This section provides an overview of the contents of each chapter, and the publications that it draws on. These are the chapters shown in black boxes in Figure 1.

Chapter two provides context and sets the scene for how economic evaluation currently includes carers' HRQoL (Objective 1). It provides an example of an economic model that only includes patients' HRQoL and demonstrates how current methods adapt this to also include carers' HRQoL. It describes the data and data sources that are typically used to populate economic models that include carers' HRQoL. This chapter discusses the limitations of both the data and the methods and highlights the relevance and importance of the following chapters. This draws on my findings from Pennington et al (2022) and Pennington & Al-Janabi (2024) as well as work undertaken prior to this PhD (Pennington, 2020).

Chapter three describes a review of studies that explore how carer health outcomes/wellbeing change over time (Objective 2). This was inspired by the evidence gap for longitudinal evidence for carers' HRQoL identified in Chapter two and was designed to inform the analysis undertaken in Chapter four. I found that there was relatively little published research on longitudinal studies of carers' HRQoL, so I broadened the scope to also consider methods that were used to analyse longitudinal evidence on carers' health or wellbeing, and variables that were included in cross-sectional analysis of carers' HRQoL.

Chapter four describes the estimation of carer utilities and the causal effect of caring on HRQoL (Objective 3). This comprises two analyses. The first analysis uses a longitudinal dataset for within-household dyads where one person reports providing care for the other and so changes in carers' HRQoL can be considered as a function of changes in patients' HRQoL and the intensity of caring (Pennington et al, 2025b). The other analysis uses a cross-sectional survey of people's HRQoL and uses causal inference methods to estimate the causal effect of providing different weekly volumes of unpaid care (unpublished at present, included in an Appendix in Chapter 9).

Chapter five discusses methods for including carers' HRQoL in economic evaluation (Objective 4). This builds on chapter two by visiting the theories underlying care research and discussing how these relate to current methods. The relationship between bereavement and caring is discussed, because of the importance of the implications of current methods on how patient death affects carers' HRQoL. The theory, combined with the concepts underpinning the analyses in chapter 4, are used in developing a new approach for modelling the effect of caring on HRQoL in economic evaluation. The implications of the new approach are discussed, including consideration of displaced carers' HRQoL. The new approach for modelling caregiver's HRQoL (Pennington et al, 2026) is central to this chapter, which also draws on the evidence from my analysis of the effect of bereavement (Pennington et al, 2025a), and challenges described in co-authored papers (Henry et al, 2024; Mott et al, 2023).

Finally, **chapter six** discusses the research project in its entirety, and provides a conclusion.

Figure 2 provides a visual summary of the chapters of the thesis, their relationship to each other, and their connection via the concept (illustrated by a thought cloud) that the relationship between caring and HRQoL is complex and can have negative and positive impacts. This emerged as an important theme from the background (Chapter 2) and literature review (Chapter 3) and underpins the analysis in Chapter 4 and the modelling methods in Chapter 5.

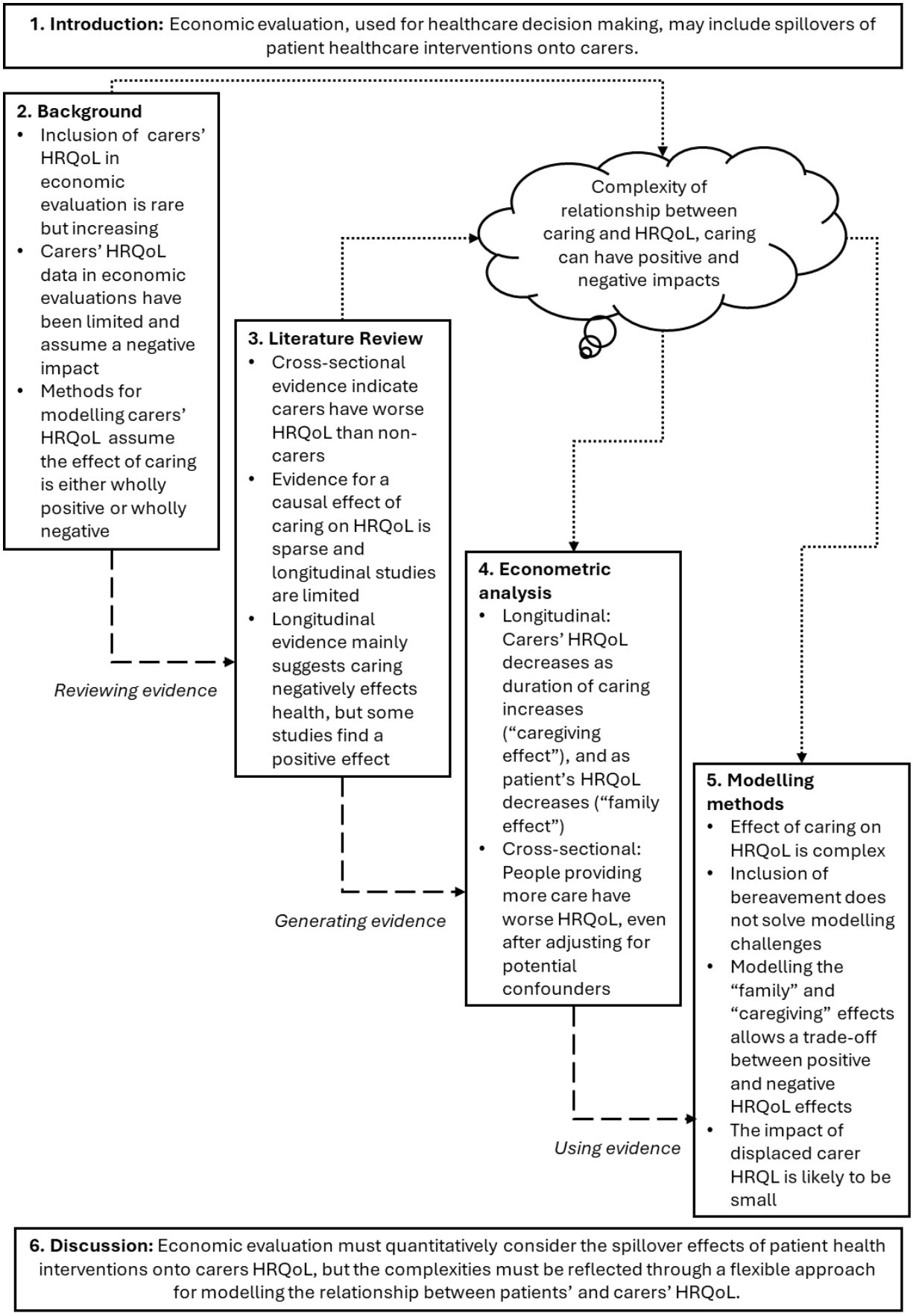


Figure 2: Visual summary of thesis

2 Background

2.1 Introduction

This chapter provides the motivation for the thesis, describing the current state of play for including carers' HRQoL in economic evaluation and the associated challenges. It introduces the concept that while carers may have worse HRQoL than non-carers, this does not prove a causal effect of caring on HRQoL, due to potential selection bias (the factors which determine whether somebody becomes a carer also influence their HRQoL). To understand whether and how caring may affect HRQoL, it draws on theoretical concepts and published analyses.

2.2 Economic evaluation

Economic models are often required to predict long-term costs and QALYs. These models simulate disease progression for a hypothetical cohort of patients with and without the intervention. They assign costs (including for the interventions and disease management) and utilities to the disease health states. A simplified example is illustrated in Figure 3, this depicts a disease that can be classified into early and advanced stages, and death. As well as the costs of the intervention or comparator (£10,000 for intervention and £2,000 for comparator), each health state has an associated cost for disease management: £1,000 for the early stage and a higher cost of £5,000 for the advanced stage, and no cost for death. Each state also has an associated patient utility (treatment-independent): 0.6 for early stage (worse than the general population who do not have the disease) and 0.4 for advanced stage (worse than in the early stage) and 0 for death. Clinical effectiveness data can be analysed and extrapolated to predict the time spent in each health state for patients receiving the intervention and comparator. The health state occupancy time can be multiplied by the associated costs and utility for each health state, to give the overall costs and QALYs for intervention and comparator.

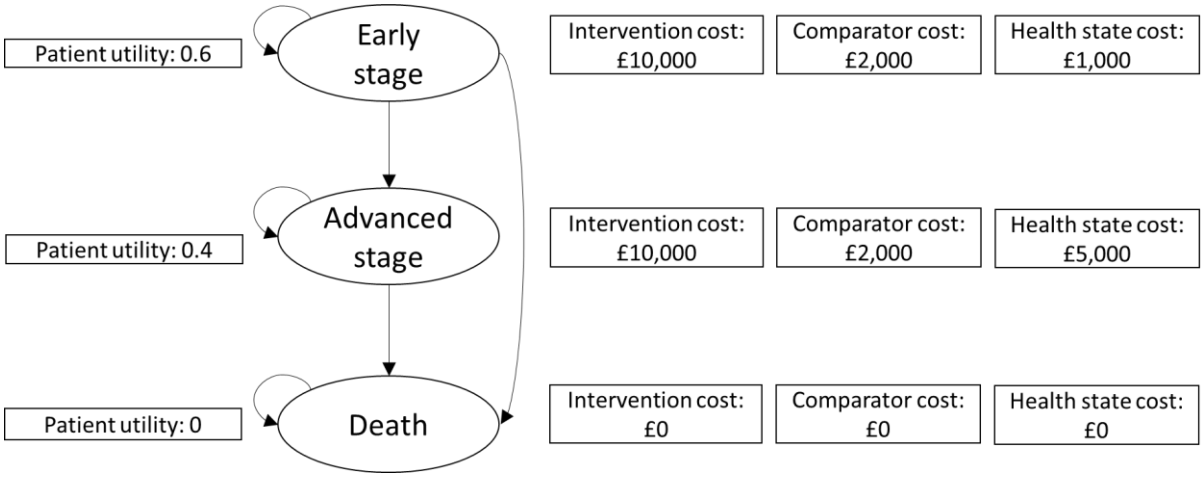


Figure 3: Illustrative economic model with only patient outcomes

2.3 Patients' and carers' outcomes

In its reference case for health technology evaluations, NICE considers that the perspective on outcomes should be "all health effects, whether for patients or, when relevant, carers" and its perspective on costs should be "National Health Service and Personal Social Services" (National Institute for Health and Care Excellence, 2022). This means that where care is provided by formal (paid) carers, the costs of this care can be included. By contrast, unpaid care is included through its effects on health, rather than through costs. NICE states a preference for cost-utility analysis, with health effects measured in QALYs. Within a cost-utility analysis framework, including QALYs for carers as well as patients is necessary and consistent with the aims of maximising the QALY gain for the population.

While health economic evaluation has historically focussed on health outcomes for the patient only, there is an increasing interest in also including health outcomes for unpaid carers. When I reviewed the inclusion of carers' QALYs in NICE Technology Appraisals (TAs) and Highly Specialised Technologies (HSTs) in 2019 (prior to beginning this PhD), I found that 3% of TAs and 50% of HSTs included carers' QALYs (Pennington, 2020). In a 2024 update of my review, Kanters et al found that these proportions had increased to 6% of TAs and 78% of HSTs (Kanters et al, 2024).

NICE is not the only HTA body that permits or recommends the inclusion of carers' QALYs. I analysed guidance from 13 countries and found that while two (Australia and New Zealand) explicitly preferred to exclude carers' QALYs, inclusion of carers' QALYs was permitted in the reference case for four countries (England, France, Ireland and the Netherlands), depended on specific circumstances in three countries (Canada, German and the US), and was unclear for the remaining four (Wales, Scotland, Sweden and Italy) (Pennington et al, 2022). To better understand how carers' QALYs had been included in economic evaluation for the

different HTA bodies that permit this, I chose five recent case studies for technologies in different disease areas, where I knew that NICE had included carers' QALYs. Choosing case studies where NICE had included carers' QALYs meant that I knew at least some HTA bodies would include carers' QALYs, so was a pragmatic decision. I explored what the other HTA bodies had done in their appraisal of these technologies. I found that the different HTA bodies tended to use the same sources of carers' HRQoL data often without considering transferability between populations (different disease areas) or countries. Furthermore, the data were all cross-sectional in design, which I had previously noted was a concern when attempting to understand the causal effect of caring on HRQoL, since people who become carers may differ from people who do not become carers and those differences may also influence HRQoL (see discussion of selection bias in Section 2.5) (Pennington, 2020). Each of the examples estimated carers' QALYs using either the "absolute utility" approach (including carers' QALYs as long as the patient is alive) or the "disutility" approach (including carer QALY losses as long as the patient is alive), which I previously noted lead to very different impacts on cost-effectiveness results and have different interpretations of how we consider unpaid carers (Pennington, 2020).

An example of including carer utilities or disutilities in the illustrative model shown in Figure 3 is depicted in Figure 4. The costs are unaffected (since we are considering only health effects to carers) and so for simplicity are not presented. Carer utilities are shown in the grey boxes, and carer disutilities (relative to 1 to represent full health) in the black boxes. Using the utility approach, carer utilities are higher in the early stage (0.8) than in the advanced stage (0.7) and so delaying disease progression increases the QALY gain for carers. Using the disutility approach, carer disutilities are smaller in the early stage (-0.2) than in the advanced stage (-0.3), and so delaying disease progression decreases the QALY loss for carers. The incremental effect of delaying disease progression is identical between the two approaches. The effect of extending patient survival is more complex: the utility approach has typically assumed that the carer utility in the patient death state is 0 (carers do not gain QALYs after patients die) and the disutility approach has assumed that the carer disutility approach in the patient death state is 0 (carers do not lose QALYs after patients die), and so the two approaches lead to different results and have different interpretations. For life-extending treatments, the utility approach assumes that these always lead to a QALY gain for carers, whereas the disutility approach assumes they always lead to a QALY loss for carers. The two approaches are discussed further in this chapter (Section 2.7) and in Chapter 5.

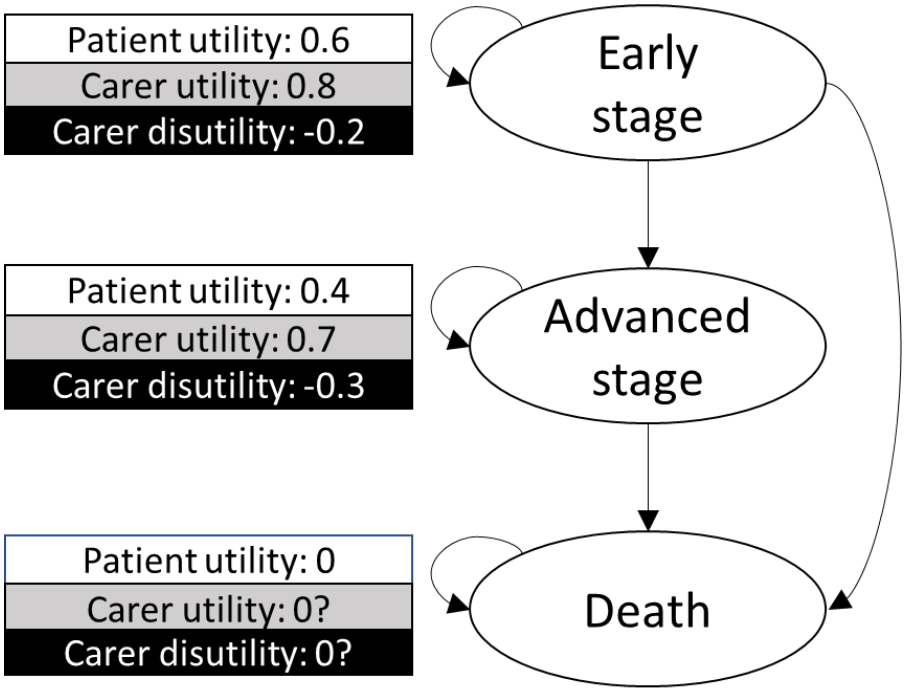


Figure 4: Illustrative economic model with patient and carer outcomes

Following my review of international guidance and examination of case studies, I made six recommendations regarding the inclusion of carers' HRQoL in economic evaluation for HTA (Pennington et al, 2022). I felt that some of the issues around methods used to model carers' HRQoL in economic evaluation warranted further exploration, and so I discussed these in my 2024 publication "Modelling Informal Carer's Health-Related Quality of Life: Challenges in Economic Evaluation" (Pennington & Al-Janabi, 2024). To some extent, some of the key points from my 2022 (Pennington et al, 2022) and 2024 (Pennington & Al-Janabi, 2024) papers overlap, and some overlap with recommendations in Recommendations for Emerging Good Practice and Future Research report published by the Spillovers in Health Economic Evaluation and Research (SHEER) Taskforce (Henry et al, 2024). I therefore summarise the key issues by topic area in the following sections of this chapter.

2.4 Measuring and valuing HRQoL

HRQoL can be measured using generic or condition-specific measures. Generic measures include the EQ-5D, the Short-Form 6 Dimension (SF-6D) and Health Utilities Index (HUI). NICE states a preference for EQ-5D to measure the HRQoL of adults (National Institute for Health and Care Excellence, 2022). The EQ-5D is a short questionnaire instrument designed to measure health-related quality of life, by asking respondents about how their health impacts five domains: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. There are two versions: EQ-5D-3L and EQ-5D-5L. Both measures contain the same domains. In 3L, there are three levels to describe problems within each

domain (no problems, some problems, extreme problems) and in 5L there are five levels (no problems, slight problems, moderate problems, severe problems, extreme problems). NICE currently recommends the 3L value set should be used and not the existing 5L value set published by Devlin et al 2018 (National Institute for Health and Care Excellence, 2019b). There are specific measures for carers' HRQoL, such as the CarerQol-7D, the Carer Experience Scale (CES) and the Adult Social Care Outcomes Toolkit (ASCOT)-Carer (McLoughlin et al, 2020), but these are not recommended by NICE. The EQ Health and Wellbeing (EQ-HWB) instrument is designed for the assessment of patients and carers and social care users, but its valuation is still in development (EuroQol, 2025).

2.5 Data to inform carers' HRQoL

Data on carer's HRQoL has typically been sourced from a small number of cross-sectional studies that report utilities for carers providing different levels of care. Evidence has demonstrated that carers have worse HRQoL than non-carers, and so it has been assumed that caring detrimentally affects HRQoL. Disutilities for caring have been calculated by comparing these utilities to either full health (utility value of 1) or age- and sex-adjusted expected population utilities. For example, at least 11 NICE appraisals (Pennington, 2020) and numerous international HTAs (National Institute for Health and Care Excellence, 2018a; 2019a; Tandvårds- och läkemedelsförmånsverket, 2020) used data from a cross-sectional study of HUI scores of 679 carers of patients with Alzheimer's Disease, with scores ranging from 0.86 to 0.94 (Neumann et al, 2000). The appraisals incorporated disutilities for different patient health states up to 0.14 – presumably calculated as the difference between the lowest carer utility and full health. This assumes that a carer would have HRQoL equivalent to full health if they were not caring for a patient, ignoring firstly any other impacts on HRQoL that may be expected (for example common comorbidities related to age), and secondly ignoring any differences between people who become carers and those who don't. In 2006, when this study was first identified, the evidence review group who identified this study noted a number of limitations with it, including its cross-sectional design (Loveman, 2006). Almost twenty years later, this study is still being treated as evidence that improving a patient's health will improve carer's HRQoL.

Using such data in economic models to predict carers' HRQoL from patient health status assumes that as the health of the person they care for decreases, the carer's HRQoL necessarily worsens. However, this ignores selection bias and assumes that all differences in different carers' (including non-carers') HRQoL are causal effects of caring. It also assumes that patient health status is the only mechanism which affects carers' HRQoL, when evidence has demonstrated that there are at least six mechanisms by which interventions may affect carers' HRQoL (Al-Janabi et al, 2019). To understand carers'

HRQoL trajectories would involve collecting longitudinal data and analysing it in the context of changes to the health of, and care (both unpaid and formal health and care services/interventions) provided to, the care-recipient.

Using data from different interventions and disease areas further compounds this problem and oversimplifies the relationship between caring and HRQoL. Furthermore, using evidence from different countries ignores the differences in the cultural and social environment and in formal care provision.

2.6 How does caring affect HRQoL?

Where economic evaluation has included carers' HRQoL outcomes, it has typically been assumed that caring detrimentally affects HRQoL: carers have worse HRQoL than non-carers HRQoL and HRQoL worsens as the volume of care provided increases. This belief has its roots in the "life course and stress process framework", and the concept that transitions, such as the transition into caring determine life course trajectories. Pearlin describes that transition into the caregiving role, and the associated related hardships, can become so engulfing that it reshapes the life course, adversely affecting the caregiver's health and wellbeing. He argues that this transition is particularly impactful since it is often unexpected and so the caregiver has not prepared for it (Pearlin, 2010).

There is evidence that caring has a positive impact on carers, and Roth and colleagues suggest that studies of population-based samples with robust methods to address confounding actually report improved health outcomes for carers (Roth et al, 2015). The "healthy caregiver hypothesis" proposes that positive outcomes are explained by healthier individuals becoming caregivers, but evidence to support this is again affected by confounding and Brown and Brown refute this (Brown & Brown, 2014). Schulz acknowledges that while carers are potentially at risk for these adverse impacts, they may also find benefit in caregiving and in helping care recipients, and notes variation within individual carers. However, Schulz summarises the overall body of evidence for negative effects of caregiving is much bigger than that of positive effects (Schulz et al, 2020).

Brown and Brown conclude that neither the caregiving stress process nor the healthy caregiver hypothesis adequately explain why caring can be associated with both positive and negative effects and argue that new theories are needed to explain this (Brown & Brown, 2014).

I have therefore looked at concepts that try to disentangle caregiving into positive and negative impacts, to underpin analysis of data and modelling approaches.

2.6.1 "Caring for" (the family effect) and "caring about" (the caregiving effect)

To better understand the mechanisms through which caring may affect HRQoL, it is helpful to consider the four phases of care (Fisher, 1990). These are:

1. Caring about: perceiving the needs in others.
2. Caring for: assuming responsibility to meet the needs perceived, and making care happen. Note that this may include organising and paying for care provided by others.
3. Caregiving: actually meeting the needs of the other person by performing caregiving tasks.
4. Care receiving: the response of the person receiving the caregiving. The response will depend on whether needs have been met and requires the attention of the person providing care.

Each of these four phases could be expected to affect the HRQoL of the carer, and, arguably, the HRQoL of the care-recipient. The care-recipient is not the focus here as their HRQoL is assumed to be captured within the economic model already, but we will revisit in Chapter 5 the implications of omitting this.

In the context of HRQoL, the phases and mechanisms by which caregiving affects the carers' HRQoL have been simplified into two effects: the family effect (or "caring about") and the caregiving effect (or "caring about"), as studied in two key publications by Bobinac and colleagues (Bobinac et al, 2010; 2011). These are discussed below.

2.6.1.1 Bobinac et al (2010)

Bobinac et al (2010) argued that the well-being effects on the significant others of patients (termed "spillover effects") may originate from two different sources: the caregiving effect and the family effect. The caregiving effect refers to the impact on a person's wellbeing of providing informal (unpaid) care to a patient, whereas the family effect refers to the fact that we care about other people and their health, even when we are not directly looking after them. These two terms and their definitions are important concepts and appear frequently in the literature, and this study has been cited over 280 times (Google Scholar, January 2025). The authors analysed the well-being of a cross-sectional sample of Dutch informal caregivers, using data from 1468 caregivers and 834 care recipients. This was reduced to 595 observations (1190 questionnaires) after deleting observations missing a variable and including only the most common form of informal caregiving (caregiving provided by household/family members). Caregiver well-being was assessed using a self-reported happiness scale ranging from 0 (completely unhappy) to 100 (completely happy).

The authors' findings confirmed their hypothesis that the family effect would be positive (when patient health improves, carer wellbeing improves) but that the caregiving effect would be negative (when caregiver burden increases, carer wellbeing worsens). There were statistically significant associations between caregiver wellbeing and their health, age, income, education level, and whether the patient had mental illness.

2.6.1.2 Bobinac et al (2011)

Bobinac et al (2011) reported a similar analysis to Bobinac et al (2010), but instead of using a self-reported happiness scale to measure caregiver wellbeing, they used the EuroQol-VAS to measure the health of the caregiver. They also used the EuroQol-VAS instead of the EQ-5D to measure patient health. They used the same data source with 1468 caregivers and 834 care recipients, reduced to 751 pairs due to missing variables. Unlike Bobinac (2010), they did not limit the sample to the most common form of caregiving and instead included a variable to reflect whether the caregiver and care recipient were blood relatives.

In the first model, the authors found a statistically significant positive relationship between patient and caregiver health (family effect) even when controlling for other determinants of health, and a statistically significant negative caregiving effect.

There was a statistically significant relationship between caregiver health and caregiver age, and between caregiver health and household income, but not between caregiver health and other variables.

2.6.1.3 Interpretation

The two studies by Bobinac and colleagues (2010 and 2011) demonstrate that carers' wellbeing and health are influenced by both the HRQoL/health of the patient they care for, and the caregiving tasks performed. In both studies, the authors highlight that there are clear implications for the practice of economic evaluations, but that the analysis is limited by the use of cross-sectional data. They note they cannot claim a causal relationship between variables because causality needs to be confirmed in longitudinal studies.

If data on carers' HRQoL is to be used in economic models that primarily evaluate changes in patients' health, there is a particular need to understand how carers' HRQoL changes as the health (or HRQoL) of the person they care for changes.

2.6.2 Mechanisms behind carers' wellbeing effects

Al-Janabi and colleagues conducted a qualitative study with carers and proposed that there are six main mechanisms through which the health and social care systems can affect carers' wellbeing (Al-Janabi et al, 2019). These are:

1. Provision of information, about patient health conditions, systems, and how to effectively care for the patient.
2. Management of care, and how this is shared between formal provision of care and informal/unpaid family and friends.
3. Patient health outcomes, including symptoms, independence, and side effects of treatment.
4. Alienation or inclusion, and how service delivery supports carers.
5. Patient compliance or engagement with care services, and the consequences this has on carers.
6. Timing and location of services for patients and contact with healthcare professionals.

This suggests that the effects of caring on HRQoL may be broader than the family and caregiving effect, but there is substantial overlap, and some mechanisms may be difficult to measure and to capture within the domains of HRQoL.

The caregiving effect closely aligns with the management of care mechanism: where informal carers need to provide more care, this is likely to affect their HRQoL. The caregiving effect may also capture some of the "patient health outcomes" mechanism, particularly where patient outcomes improve because they are more independent and therefore require less care. However, the family effect also captures some of the spillover from patient health outcomes, particularly when we consider the emotional and mental health effects of caring about a loved one. It is understandable that the caregiving effect is overestimated when the family effect is not specifically accounted for (Bobinac et al, 2011).

2.7 Methods for modelling carers' HRQoL

There is relatively little guidance as to how carers' HRQoL should be modelled, either in my own publications or the broader literature. I highlighted that modelling carers' HRQoL as a function of patient HRQoL alone may not be sufficient to model carers' HRQoL trajectories but did not propose a solution (Pennington & Al-Janabi, 2024). Similarly, I recommended explicitly stating and explaining the assumptions implicit in the method used to model carers' HRQoL (Pennington et al, 2022) but did not express a preference for any particular method.

The two most common approaches used historically in economic evaluation for HTA are the "(absolute) utility" approach where carer's QALYs are included until the patient dies, and the "disutility" approach where carers' QALY losses are included until the patient dies (see Figure 4 and Section 2.3). In each case, carer utilities or disutilities vary with patient health states, and the carer utility or disutility is multiplied by the patient life years within each health state. The "(absolute) utility" approach leads to a QALY gain for carers in all scenarios in

which patient survival is improved (assuming carer utilities are non-negative), and has been critiqued as suggesting that society values the lives of carers while the person they care for is alive, but places no value on the lives of bereaved caregivers(National Institute for Health and Care Excellence, 2021). By contrast, the “disutility” approach assumes that carers experience a QALY gain in all scenarios in which their caregiving burden is relieved, including when patients die prematurely, which has been labelled “perverse” by patient experts (National Institute for Health and Care Excellence, 2021). The result of the “disutility” approach that extending patient survival can lead to a QALY loss for carers has been termed the “the carer QALY trap” (Mott et al, 2023), due to its parallels with the “classic” QALY trap (Ubel et al, 2000): under the “classic” QALY trap interventions which extend the survival of people with poorer HRQoL are less valuable than those that extend the survival of people with higher utilities. Under the carer QALY trap, interventions which extend the lives of people who require carers are less valuable than those which extend the lives of people who do not need carers.

More recently, the “incremental utility” or “increments” approach has been suggested as an alternative approach to modelling the effect of patient interventions on carers’ QALYs (Grimm et al, 2022). This approach was preferred by the NICE committee in the appraisal of lecanemab for treating mild cognitive impairment or mild dementia caused by Alzheimer’s disease (National Institute for Health and Care Excellence., 2025). This approach involves calculating utility increments for each health state relative to some baseline, typically the worst alive health state. In the example in Figure 4, if the baseline health state was “advanced stage” then there would be no utility increment in that health state, and the utility increment in the “early stage” state would be 0.1 (0.8 – 0.7). This means that extending survival in the “early stage” state would lead to a QALY gain for carers, but this is much smaller than it would be using the “absolute utility” approach. Extending survival in the “advanced stage” state would have no impact (positive or negative) on carers’ QALYs since the utility increment for this state is 0 (0.7 – 0.7).

Another alternative, proposed in the literature but not commonly used in HTA, is the use of a “multiplier” (Al-Janabi et al, 2016b). Here, the gain in patient QALYs is multiplied by a ratio of carer QALY gains to patient QALY gains. In the example in the paper describing this approach, the ratio is 0.16, calculated from the coefficient in a regression model to estimate family members’ EQ-5D as a function of patient EQ-5D (Al-Janabi et al, 2016c). This approach assumes a uniform relationship between patient QALY gain and carer QALY gain (the carer QALY gain is not influenced by how the patient QALY gain was derived, whether it is from mortality or morbidity improvements). The authors note that if the multiplier is constant across interventions, it will make no practical difference to decision-making, since

the QALY gain for all interventions would increase by the same amount, and the displaced QALY gain and hence threshold would also decrease by the same amount. The Dental and Pharmaceutical Benefits Agency in Sweden recently proposed a similar method, in which carers are assumed to gain a proportion of the patient's QALY gain during the period the patient would have lived with standard care. They term this proportion the "standard rate" and note that the magnitude of this is not currently informed by data and needs to be determined (The Dental and Pharmaceutical Benefits Agency (TLV). 2022).

Even once carers' QALYs have been calculated, aggregation of patient and carers' QALYs is not straightforward: it is unclear whether they are equally valued (Al-Janabi et al, 2022) and interchangeable. Guidelines recommend that they are reported in disaggregated and aggregated format.

2.8 Conclusion

The inclusion of carers' HRQoL in economic evaluation is increasingly common but still poses challenges for analysts and decision-makers. These challenges are well described in the literature, but clear guidance does not yet exist to address them. This thesis aims to address the challenges related to data to inform carers' HRQoL in economic evaluation, and methods to model carers' QALYs.

3 Evidence to inform the effect of caring on health-related quality of life

3.1 Introduction

Chapter 2 described that economic evaluation may include HRQoL, measured by utilities, for carers as well as patients, and described some of the data that has typically been used to inform carers' utilities. As discussed there and in my 2022 and 2024 papers (Pennington et al, 2022; Pennington et al, 2024), much of the evidence that has been used to estimate carers' QALYs in economic models stems from cross-sectional data. This measures carers' HRQoL at one point in time and can be used to compare carers' HRQoL to that of non-carers, or between carers providing different levels of caregiving. However, this does not account for selection bias: people who become carers may be fundamentally different to people who do not become carers and their differences in HRQoL may not be entirely due to caregiving. For example, older people may be more likely to become carers than younger people and older people may have worse HRQoL than younger people, so the difference between carers' and non-carers' HRQoL may be due to ageing as well as caring. Similarly, having lower levels of psychological wellbeing and higher levels of stress or being unemployed may be associated with both lower HRQoL and being more likely to become a carer, so these factors may partly explain some of the differences between HRQoL of carers and non-carers (Brown & Brown, 2014). Ignoring selection bias assumes that all the differences between carers' and non-carers' HRQoL is entirely due to caregiving. Using these values in economic models would imply that when someone becomes a carer (or stops being a carer), all these other factors also change. This may overestimate carers' HRQoL effects and bias the results. Furthermore, there may be selection into the amount of caregiving: carers who provide more care may differ from those who provide less care.

I highlighted that there is a need to analyse longitudinal data to mitigate the effect of selection bias in analysing the effect of caregiving on HRQoL. Longitudinal data can help us understand how people's HRQoL changes as they become a carer, their caregiving changes, or they stop providing care (Pennington et al, 2022; Pennington et al, 2024). This was echoed as a research recommendation by the Spillovers in Health Economic Evaluation and Research (SHEER)Taskforce (Henry et al, 2024). Scholars researching the theories behind the effects of caregiving have similarly noted the limitations of the current data and the paucity of longitudinal evidence and well matched carer and non-carer samples (Brown & Brown, 2014; Roth et al, 2015). This section reviews the literature to understand the existing evidence base for longitudinal data on changes in carers' HRQoL over time. This can then inform new analyses and modelling methods in later chapters.

3.2 Aim

My literature review aimed to answer the research question “How does carers’ HRQoL change as the health (or HRQoL) of the person they care for changes?” The purpose of the review was to provide context for my research, identify prominent studies, and inform and inspire the analysis that I would undertake. The review was a narrative review and did not aim to comprehensively identify all relevant studies, and so a full systematic literature review was neither necessary nor appropriate. However, I followed a systematic approach, aiming to identify relevant papers that would inform the state of research in this area to date.

I wanted to identify longitudinal studies that reported patients’ and carers’ HRQoL at multiple time points and analysed the relationship between changes in patients’ and carers’ HRQoL. However, I recognised that there would be few studies of this nature, and so I broadened the scope of the review to inform the methods I could use by looking for longitudinal studies that analysed how carers’ health or wellbeing (rather than only HRQoL) changed as patient health or caring responsibilities changed.

I also drew on three sequentially updated published systematic reviews of carers’ HRQoL. These provided an overview of the evidence base for (primarily cross-sectional) studies reporting carers’ HRQoL and were useful in understanding the limitations of currently published studies and evidence gaps.

This chapter discusses the evidence to inform:

- Predictors of carers’ HRQoL: Section 3.4.2
- Predictors of changes in carers’ HRQoL: Section 3.4.3
- Predictors of changes in carers’ health or wellbeing: Section 3.4.4

3.3 Method

It is well understood that MeSH terms and free text searching are problematic in the area of HRQoL (utilities). This is in part due to the definition of HRQoL, that cost utility analyses use this term to refer specifically to utility measures that anchor HRQoL between 0 and 1 to represent death and full-health, whereas the wider literature may use this term to describe much broader measures and concepts. There are no dedicated thesauri terms for commonly used HRQoL measures, and the broader terms of “quality of life” are generic and while they maximise sensitivity, are not very precise (Papaioannou et al, 2013).

Reviews of cost utility analyses and other literature reviews within healthcare tend to focus on searching only biomedical literature, using databases such as MEDLine via PubMed. The papers that I had already identified as being relevant (through my background research and source data in NICE appraisals) were indexed in PubMed, so this represented a good

source of information, specific to HRQoL data. However, I was aware that broader social sciences literature may also be informative, particularly in the case of analysis of panel datasets. This literature, in addition to the biomedical literature, would provide a richer source of information, but would be less specific to health – for example also covering labour outcomes.

“Pearl-growing” and “berry-picking” are approaches that may be used to identify studies where full systematic reviews are not required. The concept behind pearl growing is to identify one (or more) particularly important study initially and use the information from that study to grow the pool of papers. This is done by identifying free-text and thesaurus terms on which to conduct a specific search (Booth, 2016). Given the challenges noted above in searching for HRQoL studies, I determined that pearl-growing was unlikely to be helpful in this review. Berry-picking refers to the idea of maximising the amount of relevant literature found by focussing reviewing in one “patch”, before moving on to another “patch” – the aim being to identify key studies rather than to systematically identify every study (Bates, 1989). Berry picking has six main elements:

1. Following up references
2. Forward citation searching
3. Journal searching
4. Area scanning
5. Subject searching in databases
6. Author searching

I used an iterative search strategy, using two key papers as my starting point, and following some of the berry-picking techniques until I reached saturation (finding no new relevant papers).

The two starting papers were the only two studies I was aware of that had assessed change in patient and carers’ HRQoL simultaneously – I had identified these through discussion with other researchers working in this area. One study aimed to predict changes in the HRQoL of family members of a person with meningitis (Al-Janabi et al, 2017) and the other compared the validity and responsiveness of five measures of HRQoL in carers of patients across multiple conditions (McLoughlin et al, 2020).

Initially, I performed forward citation searching using MEDLine via PubMed (as an initial focussed search to identify HRQoL papers) and followed up the references in the papers to identify new studies. I then categorised the identified studies by whether they reported each of the following:

1. Cross-sectional analysis of carers' HRQoL,
2. Cross-sectional analysis of carers' health, but not specifically HRQoL
3. Longitudinal analysis of / predictors of change in carers' HRQoL
4. Longitudinal analysis of / predictors of change in carers' health, but not specifically HRQoL

I added studies which fell into categories 1, 3 and 4 to my "pool" of studies and followed up their references and forward citations and repeated this process. (I did not add studies that predicted carers' health but not HRQoL to this pool (category 2) since it would be a large number of potentially irrelevant studies (due to their cross-sectional design, methods that do not account for selection bias, and lack of overlap in outcomes) and I wanted to focus on the most relevant studies). This returned relatively few studies, suggesting a point of saturation.

To identify papers from outside the medical literature and ensure that my search was as comprehensive as possible, I selected two key studies which were frequently cited by identified papers for forward citation searching using Google Scholar. Google Scholar searches across disciplines and sources, and is not limited to published studies, also including for example theses and conference proceedings. These two key studies were the 2010 and 2011 studies by Bobinac and colleagues and are pivotal in informing the concepts underpinning research into carer's health (Bobinac et al, 2010; 2011). These studies decomposed the effects of care into the "family effect" and the "caregiving effect" (see Section 2.6) and so I was interested in studies that captured spillovers on either close family members or informal caregivers (recognising there will be some overlap). I therefore included studies where the exposure was either informal caring or being a family member of a person with a health condition (or both), with the aim of capturing both the family effect and the caregiving effect. I excluded studies where the exposure was an intervention targeted at existing carers.

I excluded studies which specifically studied the effect of caregiving due to covid, since I was interested in the effect of caregiving generally and not limited to one specific event or time period which may not be generalisable given all of the changes imposed on society and individuals during the covid pandemic.

I also reviewed papers included in a themed issue of Pharmacoeconomics which focussed on measuring family spillover effects of illness (Vol 37 Issue 4 April 2019).

3.4 Results

3.4.1 Summary

While there is a growing body of evidence providing HRQoL estimates for carers, studies typically only report the average utility for carers in a specific population (for example caring for patients with a specific health condition). Studies are typically cross-sectional in nature, and non-comparative. Where comparisons are made to estimate disutilities, they often compare to population averages rather than well-matched samples. As such, while the body of evidence suggests that carers experience worse HRQoL than non-carers, it does not robustly demonstrate that caring causally affects HRQoL. The need for longitudinal data to capture the impacts of caring on HRQoL is noted in the literature (Theakston & Mott, 2025).

There is a paucity of evidence estimating a causal effect of caring on HRQoL using longitudinal studies. To date, only a handful of studies have examined change in carers' HRQoL, and this has been limited to 2-4 time points, a maximum of two years' apart. These studies all suggested that carers' HRQoL is affected by changes in patient health but are somewhat limited by their design and duration.

There is much more longitudinal evidence analysing the effect of caring on broader wellbeing or mental health outcomes, as demonstrated by the existence of two systematic reviews and 21 additional studies. These studies used a range of different methods and datasets but consistently found that (transitioning into) caregiving was associated with worse outcomes for the caregiver (with few exceptions reporting no significant differences). However, three studies sought to separate the caregiving effect from the family effect: one of these found that both effects were negative, and the other two suggested that the caregiving effect can alleviate the negative family effect (in some cases). This highlights the importance of controlling for the family effect when attempting to isolate and analyse the caregiving effect. Taken together, these results also suggest that while worsening care-recipient health is consistently associated with worse caregiver outcomes, the caregiving effect may act in different directions for different situations.

3.4.2 Predictors of carers' HRQoL

3.4.2.1 Summary

Typically, carers' HRQoL is predicted only by the presence or absence of the patient's health condition, although in some cases severity may also be included. These studies provide evidence of a spillover effect on carers' HRQoL across a range of disease areas, but do not report change in carers' HRQoL alongside change in patient HRQoL. These are not informative for how carers' HRQoL changes.

There are two published systematic reviews by the same authors from 2013 and 2019 that report spillover effects on caregivers or family member's utility and provide context on carers' HRQoL in cross-sectional analysis (Wittenberg et al, 2019; Wittenberg & Prosser, 2013). A different team updated this review in 2024, which was presented at the Health Economists' Study Group Winter 2025 but is not yet published in a peer-reviewed journal (Theakston & Mott, 2025).

All three systematic reviews demonstrate that information on carers' HRQoL is available across a range of conditions, but that this data is seldom from well-matched comparative studies. Given the issues of selection into caregiving, it is likely that much of the data on disutilities (where it is available) is subject to bias, and that it may not accurately reflect the difference in HRQoL that is due to caregiving. The non-comparative data (such as commonly reported absolute utilities) or data that only focusses on a comparison between caregivers of patients with or without a specific disease, is not useful in an economic evaluation which may reflect an improvement or worsening of the patient's disease.

3.4.2.2 Wittenberg et al (2013)

Wittenberg et al (2013) searched the medical, psychology and economics literature from inception to February 2012 to identify published studies measuring spillover on family members' and caregivers' HRQoL. They included 15 studies, seven of which reported a disutility, and eight of which reported data from which the authors calculated disutility. Ten studies reported spillover disutilities for family members or caregivers of specific patient populations, one study reported disutilities across chronic conditions (Wittenberg et al 2013 analysis of MEPS data, described in Section 3.4.3.5), and two were designed to develop measures of caregiver utility. One study directly assessed spillover disutility using a chained time trade-off method (Basu & Meltzer, 2005), all other studies estimated spillover disutility by comparing utilities from family members or caregivers to a comparison or control group. The comparison groups included population norms (n=7), empirically selected control samples (n=6), within-subject scores across time points (n=1) and within-subject hypothetical comparisons (n=2), noting some studies used two comparator groups. The authors concluded that although caregiver burden was well-recognised, evidence of spillover disutility was sparse and that their findings suggested that caregiver disutility was small but measurable. They noted that estimating spillovers from comparison to population norms may be obscured or confounded by self-selection into caregiving and that accurate and valid measurement of caregiver HRQoL is necessary for its inclusion in economic evaluation.

3.4.2.3 Wittenberg et al (2019)

Wittenberg et al (2019) searched the medical and economic literature from inception to April 2018 to identify articles reporting preference-based measures of caregiver or family HRQoL. In this updated review they found 80 studies, demonstrating that the evidence base had increased substantially over the intervening five years. The authors grouped the studies into three categories: eight studies reported spillover utilities or disutilities only, 25 reported caregiver/family utilities with a matched population or comparison group, and 50 reported caregiver/family utilities only (the authors do not explain why this sums to 83). The authors noted that comparisons to population norms could be used to derive spillovers for the 50 studies that report only caregiver/family utility, if they are carefully considered to match the demographics of the caregiver sample. The authors concluded that the scope of economic evaluations is expanding to include caregivers and families, but that this raises practical challenges and highlights data gaps, and that additional research is needed on the methods of measuring and incorporating spillover QALYs into cost utility analyses.

3.4.2.4 Theakston & Mott (2025)

Theakston and Mott (2025) replicated Wittenberg et al's 2019 search strategy and repeated it for the period from April 2018 until July 2024. They also added EQ-HWB into their search terms, as it was not developed at the time of the previous review but was anticipated to be relevant. They included 96 studies in the review, emphasizing the growth in this research area in recent years. Where studies reported one specific condition, Alzheimer's and dementia were the most commonly studied conditions (n = 30 total), followed by cancer (n=8), autism (n=7) and mental health disorders (n=7). EQ-5D was the most commonly used instrument to measure HRQoL (n=71), and CarerQol-7D was the most commonly used carer-specific measure (n=21). 88 of 89 studies reported absolute utilities, disutilities were less commonly included; the comparator was typically general population but from a range of sources. Perhaps most informatively, Theakston and Mott categorised their included studies as primary studies, secondary data analyses, trials, and direct elicitation. Within these categories, they specified that they found nine primary longitudinal studies and three longitudinal secondary data analyses. The authors kindly provided me with a list of references for the 12 studies within these categories. Of the nine primary studies, some were not longitudinal for carers (longitudinal data was only provided for patients) (Hughes et al, 2022; Mulders-Manders et al, 2018), and some only included carers and not patients (Al-Rabayah et al, 2022; Potter et al, 2023; Sukchokpanich et al, 2023; Wrotek et al, 2023). One study did report changes in patients' and carers' HRQoL and used multivariable linear regression to examine the relationships between the change scores and patient and carer characteristics. They found a small statistically significant decrease in carers' HRQoL at 24

months, and the only statistically significant relationship between change in carer's HRQoL and other variables was with carers age: older carers reported a greater reduction in their HRQoL (Park et al, 2019). One study described changes in patients' and carers' HRQoL at six, 12 and 18 months but did not analyse these changes together (in fact, carers' HRQoL changed significantly but patients' HRQoL did not) (Froelich et al, 2021). McLoughlin, described earlier, was the only other primary study that predicted changes in carers' HRQoL (McLoughlin et al, 2020). The three secondary data analyses of longitudinal data included one study that used pooled cross-sectional analysis (Lee et al, 2022), one study of the effect of loneliness on carers (Majmudar et al, 2024), and one study of the effect on carers of a family involvement programme (Musters et al, 2024).

Theakston and Mott also noted that some clinical trials measured carer's HRQoL, including a study by Tubeuf and colleagues that predicted carers' HRQoL as a function of patient HRQoL, amongst other variables (Tubeuf et al, 2019). This study presented data from a randomised controlled trial comparing family therapy with treatment as usual as an intervention for self-harming adolescents. The trial collected EQ-5D-5L for adolescents at baseline, six and 12 months, and HUI2 for one parent at the same time point. The authors modelled the utility of the parent at each time point as a function of the utility of the adolescent at that time point, controlling for other characteristics. This study was also identified through searching the PharmacoEconomics special edition.

3.4.3 Predictors of change in carers' HRQoL

3.4.3.1 Summary

There are far more studies reporting predictors of carers' HRQoL than of predictors of change in carers' HRQoL. Only four studies estimated the impact of changes in patient health on carers' HRQoL. Three analysed trial/survey data collected for this purpose (Al-Janabi et al, 2017; Bhadhuri et al, 2017; McLoughlin et al, 2020) and one analysed a routine dataset (Wittenberg et al, 2013). All included EQ-5D (3L or 5L) as a measure of carers' HRQoL and two additionally included other HRQoL measures (Bhadhuri et al, 2017; McLoughlin et al, 2020).

All four studies suggested that carers' HRQoL is affected by change in patient health, but the changes tended to be small and not always statistically significant. All four studies considered relatively short time periods, with Wittenberg et al (2013) the longest at two years. Each study only considered carers' and patients' HRQoL at two time points: baseline and follow-up, and so assumed that patient and carers' HRQoL changes were contemporaneous. It may be reasonable to assume there could be a lagged effect, whereby the carer takes some time to adjust to the patients' changed HRQoL. Furthermore, the

studies include limited data on the amount of unpaid care provided and the extent to which this complemented or replaced formal care provision, so they assumed that either formal care provision did not change or had a negligible effect. My interpretation of the findings by McLoughlin et al (2020) is that change in informal care hours may be a better predictor of changes in carers' HRQoL than changes in the patient HRQoL (but these are likely to be highly correlated). This may suggest a stronger caregiving effect than family effect, but the authors analysed the two effects separately, whereas Bobinac et al noted that analysing the caregiving effect without explicitly considering the family effect led to an overestimate of the caregiving effect (Bobinac et al, 2010; 2011).

3.4.3.2 Al-Janabi et al (2017)

Al-Janabi et al (2017) aimed to predict changes in individual carers' HRQoL from changes in patients' HRQoL (both described using EQ-5D-5L) over 12 months. Their rationale for using data from two time points rather than a cross-sectional design was to provide a more direct test of the predictive validity of models as it is ultimately changes in health status that are used in economic evaluation. They used data from a UK-wide survey of 497 carer-patient dyads of meningitis survivors and their family members. Most patients' and carers' HRQoL changed by between -0.2 and +0.2.

The authors estimated four predictive models (considering patient EQ-5D-5L items, specific consequences of meningitis, patient healthcare use, and carers' characteristics plus patient healthcare use) of change in carers' utility. In all models the authors additionally controlled for patient's age and sex. They estimated all models using linear regression, but also investigated adjusted limited dependent variable mixture models and a Beta regression model, noting that comparisons between these revealed little difference in p-values and better Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) for the linear regressions. They found the predictor variables in all four models were mostly non-significant and, in some cases, there were anomalies, for example associations between improvements in patient anxiety/depression and a decrease in carers' HRQoL. The authors proposed that the low predictive accuracy (measured by the correlation between observed and predicted change, absolute errors, mean absolute error and root mean squared error) is because patient-related variables on their own are likely to be poorly correlated with carers' HRQoL and other factors will also influence carers' HRQoL. The authors noted that predictions are likely to be better at a group level as various external influences on carer's HRQoL may cancel out.

3.4.3.3 McLoughlin et al (2020)

McLoughlin et al (2020) surveyed carers' HRQoL at baseline and 12 months follow-up using EQ-5D-5L, three carer-quality of life measures (Carer Experience Scale (CES), CarerQoL-7D, ASCOT-Carer), and a capability measure (ICECAP-A). The survey also contained contextual questions related to the carer, care recipient, and the caring situation. The authors assessed construct validity of the measures using convergent validation and assessed the relationship between the change in carer's HRQoL and changes in two anchors (external indicators of a change in a related outcome): care recipient HRQoL measured using EQ-5D-5L (reported by the carer), and informal care hours. These were subdivided into three levels depending on whether the anchor had increased, decreased, or not changed, using minimal clinically important differences of 0.063 for EQ-5D-5L (previously estimated in the literature in 2017(McClure et al., 2017)), and changing through a threshold of either 20 or 50 hours of informal care hours per week. The authors measured responsiveness of the four measures using the standardised response mean effect size values, calculated as the ratio of mean change between baseline and 12 month score to the standard deviation of the score changes. This calculation standardised the changes so that the four different measures could be compared. They classed the mean effect sizes using the classification proposed by Cohen's d, where effect sizes below 0.2 are negligible, between 0.2 and 0.5 are small, between 0.5 and 0.8 are moderate, and above 0.8 are large (Sullivan & Feinn, 2012). The authors hypothesized that a significant improvement in patient HRQoL or a significant reduction in informal care hours would be associated with an improvement in carers' HRQoL. They also assessed the gradient of effect: this recognised that HRQoL may improve or worsen for all carers, but that a bigger improvement in carers' HRQoL would be expected to correspond to an improvement in the care recipient's health or reduction in hours of care provided, and a bigger worsening in carers' HRQoL would be expected to correspond to a worsening in the care recipient's health or increase in hours of care provided

The anchor of care recipient HRQoL improved at follow-up for 94 participants, had no change for 80 recipients, and had worsened for 97 participants. The effect sizes were trivial for CarerQoL and ASCOT-Carer, and generally small for CES, EQ-5D-5L and ICECAP-A. There was an expected gradient of effect for CarerQoL, with greater improvements in the CarerQoL score where care recipient health status improved. EQ-5D-5L scores decreased for all carers, and ICE-CAP A scores improved for all carers, with no clear gradient of effect.

The anchor of hours of informal care had decreased at follow-up for 49 participants, had no change for 193 participants, and had increased for 61 participants. The effect sizes were trivial to small for all measures, but there was an expected gradient of effect (the change in

HRQoL for the two subgroups of carers as categorised by the anchors was ordered in the expected direction in relation to the change in the construct) for all five measures. The authors noted that the small effect sizes should be viewed in the context of fluctuating carers' HRQoL and wellbeing over the 12 months (due to changing access to formal care services and care tasks, as well as the underlying quality of life of the carer) and suggest that the gradient of effect may provide a better measure of responsiveness, since this was as expected for all five measures for hours of care provided.

3.4.3.4 Bhadhuri et al (2017)

Bhadhuri et al (2017) assessed the construct validity and responsiveness of EQ-5D and SF-6D for measuring the effect of patient health conditions on family members. They considered separately carers (family members who provide care) and non-caring family members, based on responses to a survey of families of meningitis survivors. The sample included 648 non-carers and 199 carers with EQ-5D and SF-6D reported at baseline and after 12 months.

The authors hypothesised that over 12 months, the change in family members' HRQoL would be positively correlated with changes in patient HRQoL scores (carers and non-carers) and negatively correlated with changes in number of hours providing informal care (carers only). They used anchor-based methods to examine whether the HRQoL scores changed in the expected direction, with anchors selected for patient HRQoL and informal care hours, categorised by whether they increased, decreased or did not change in an important way. They defined importance for HRQoL as the minimal clinically important difference of 0.074 (previously estimated in the literature, but using a different source (Walters & Brazier, 2005) to McLoughlin (McLoughlin et al, 2020)) and for carer hours as a change of more than five hours per week. They used Cohen's d to assess the effect size, using the same classification as McLoughlin et al (2020).

There was no gradient of effect to suggest that change in patient HRQoL was correlated with change in non-carers' HRQoL, using either EQ-5D or SF-6D. There was a gradient of effect in the expected direction for carer's EQ-5D (but not SF-6D) as patient health status changed, but not as hours of care provided changed.

The authors discussed that the small or unobservable gradients of effect may be due to the short time period of the study and the relatively small changes in patient HRQoL and carer hours.

3.4.3.5 Wittenberg et al (2013)

Wittenberg et al (2013) analysed data from the Medical Expenditure Panel Survey (MEPS): a set of household-based in-person interview surveys in the US. They used household component data on family medical conditions from 5 time points, and two EQ-5D

administrations, over a two-year span, from three panels beginning in 2000, 2001 and 2002. The sample consisted of 24,188 adults representing 14,043 households and including 11,867 children. EQ-5D was collected from adults only, so the analysis considered the impact of household member's conditions on adults, but the people with conditions could be adults or children since medical condition data were collected from both. They did not specifically analyse informal/unpaid carers, but looked at spillovers to household members.

The authors used multivariable regression analysis using general estimating equations (GEE) to estimate the independent spillover effect of household members' chronic health conditions on an adult's EQ-5D score. The GEE controls for clustering at the household level. They used a static model to predict the adult's EQ-5D score at the second administration as a function of household members existing and new health conditions, and a change-score model to predict change in the adult's EQ-5D score as a function of a new diagnosis of a household members' conditions. Both models additionally included other variables known to predict utility: the presence of chronic conditions, age, race, gender, marital status, household income and household composition.

In the static model, the authors modelled EQ-5D in continuous form using Duan's two-stage approach to account for the high proportion of 1 scores. The first part of the model predicted the odds ratio of having an EQ-5D score of 1 versus <1: this was statistically significantly lower for existing mental and respiratory conditions among other (child and adult) household members, existing musculoskeletal disorders among adult household members, new mental disorders among adult or child household members, and new nervous/sensory or musculoskeletal system disorders among other adult household members. The largest effect was from the adult's own existing or new health conditions (as may be expected). There were also associations from some of the other variables. The second part of the model included only adults reporting EQ-5D less than 1, and found some significant correlations, but that chronic conditions among other household members were less consistently associated with a lower EQ-5D score.

In the change-score model, the authors created an EQ-5D difference score (EQ-5D at second administration minus EQ-5D at first administration), split into 0.05 increments to create a discrete outcome with a symmetric and approximately normal distribution but with a high peak at 0 indicating no change. This showed that occurrence of only a new musculoskeletal condition among other adults in the household was significantly associated with a reduction in the adult's EQ-5D score by 0.007. The occurrence of some new conditions for the adult himself were also significantly associated with a reduction in EQ-5D. The authors reported that they tested their transformation by modelling the outcome with a

negative binomial distribution and found similar results. The authors concluded that their analysis provided empirical evidence of the existence of spillover effects, although they noted that their results were correlations not causal effects and that most of their results fell below the level of minimally important differences for utilities.

3.4.4 Predictors of changes in carer health or wellbeing

3.4.4.1 Summary

I identified studies in the literature that uses longitudinal analysis of panel data to estimate impacts of informal caregiving on other health-related outcomes or wellbeing-based outcomes. While not directly relevant to my research question about how changes in carers' HRQoL, I recognised that some of these studies may be useful for underpinning conceptual models about how carer health changes and may suggest datasets I can analyse or econometric methods that may be suitable.

Two studies were recent systematic reviews of longitudinal evidence (Bom et al, 2019; Ervin et al, 2022), and thus provided a helpful summary on the status and methods of this research to date. Bom et al (2019) reviewed the causal impact of providing care to the elderly on the health of the caregiver, and Ervin et al (2022) reviewed the longitudinal association between providing unpaid care (to any persons requiring care) and mental health.

I identified an additional 21 studies which used longitudinal data to analyse the relationship between unpaid caregiving and health or wellbeing outcomes. These are summarised in

Table 1 and discussed below.

3.4.4.2 Bom et al (2019) systematic review

Bom et al (2019) conducted a systematic review of biomedical and social sciences literature to answer the question “What causal impact does providing informal care to elderly or older family members have on the health of the caregiver? And how does this caregiving effect differ between subgroups of caregivers?” Fifteen articles met all inclusion criteria, including using one of the following methods: propensity score analysis, simultaneous equation models (instrumental variables), regression discontinuity designs, difference-in-difference (DID) models or Heckman selection models, and excluding studies using a matching design that does not match on caregiver health prior to starting caregiving, to ensure included studies considered selection into caregiving. The authors reported the health measures included: none of the studies included preference based HRQoL measures, although the related SF-36/SF-12 physical component score (SF PCS) and mental component score (SF MCS) subscales were included in 3 studies. Twelve studies used longitudinal data, although in one case data from only one wave was used, and three studies used cross-sectional data. The definition of informal caregiving differed between studies, with five studies specifying caring for a parent, one specifying caring for a spouse, five describing caring for anyone/family member/friend, one specifying caring for someone with a specific illness, and two defining caregivers as persons living with a family member/spouse in need.

With one exception, all studies reported a short-term negative effect for certain subgroups of carers. The authors proposed that the reason for the exception could be that it defined informal caregivers as members of a household where someone needs care (without determining if they provided care). Studies generally found that caregiving resulted in higher prevalence of depressive symptoms and lowered mental health scores, but the estimates of physical health impacts were more variable, and in some cases positive. The authors discussed that this appears to be related to timing, with positive short-term impacts on self-reported health and longer term impacts on pain and drug use. Important subgroups included sex and the intensity of provided care (a stronger negative effect of caregiving for women and people providing more intensive care). The authors concluded that there was evidence of negative health effects of informal caregiving for subgroups of informal carers, and that consideration of the subgroups is important.

3.4.4.3 Ervin et al (2022) systematic review

Ervin et al (2022) conducted a systematic review of biomedical and social sciences literature to answer the question “What is the longitudinal association between unpaid caregiving and the mental health of working age adults in high-income OECD countries?” Thirteen studies

met all inclusion criteria, including using one of the following methods: linear mixed effects regression, linear fixed effects regression, ordinary least squares (OLS) regression, logistic regression. The authors reported the health measures included: none were reported to include preference based HRQoL measures, although the related SF-36/SF-12 PCS and MCS subscales were included in three studies. The definition of informal caregiving differed between studies, with two studies considering family caregiving, one specifying adult caregiving, four describing caregiving, four specifying informal caregivers/caregiving, one parental caregiving, and one sandwich caregiving (simultaneously caring for children and parents).

Two studies reported no significant difference in outcomes between family caregivers and non-caregivers. One study found significant effects only for intermittent carers (more than one episode of caregiving) or long-term care for women. One study found that caregiving, only where it imposed limitations, was associated with higher depression. All other studies reported a negative association between caregiving and mental health. Some gender differences were noted, with consistent negative association between mental health outcomes and caregiving for women, but some studies finding no or positive association between mental health outcomes and caregiving for men. The authors concluded that there was overwhelming evidence that unpaid caregiving negatively affects working age adults' mental health and recommended that the disproportionate caregiving load on women is a crucial requirement for policy development.

3.4.4.4 Additional studies

The 21 additional longitudinal analyses are summarised in Table 1

Table 1.

The studies all used large national (or international) panel datasets including two UK datasets (the English Longitudinal Study of Ageing and the UK Household Longitudinal Survey).

None of the studies' outcomes included HRQoL in a form that can be used in cost-utility analysis. Instead, they mainly focussed on measures of wellbeing or life satisfaction and mental health (including depression), with fewer studies including physical health.

Three studies included some measure of both the caregiving effect and the family effect (Bom et al, 2019; Han, 2023; Han et al, 2021). Two studies included only the family effect (Henry & Cullinan, 2024; Jain & Ma, 2024) and did not record whether the family member provided unpaid care. Sixteen studies included only the caregiving effect.

The studies that included only the family effect concluded that the presence of health conditions or difficulties in one family member have spillover effects on the depression of other family members, but this was not always significant at the 5% level. These studies used generalised method of moments (GMM) dynamic models to account for the family members' health at previous time points.

The studies that included only the caregiving effect tended to conclude that informal caregiving negatively affected outcomes for carers, particularly mental health or wellbeing. The effects on physical health were less clear: one study found that informal care provision improved self-rated health and reduced activity limitations (Lamsal et al, 2024) and another found transition into care improved allostatic load (representing "wear and tear" or the impacts of distress at a molecular and cellular level(McEwen & Stellar, 1993)) for men but not women (Bertogg, 2025). Two studies found that caring improved physical health as measured by SF-12 PCS (Bom & Stockel, 2021; Stöckel & Bom, 2022) but this may be explained by the SF-12 scoring system where worse mental health scores necessarily improve the PCS score. Many of these studies used fixed-effects models (Cheng et al, 2023; Kramer & Bleidorn, 2024; Labbas & Stanfors, 2023; van den Berg et al, 2008) (in some cases asymmetrical fixed effects (Bertogg, 2025), to allow for differences between transitioning into caregiving and transitioning out of caregiving). Three used instrumental variables (Costa-Font & Vilaplana-Prieto, 2022; Eibich, 2023; Le & Ibuka, 2023), one used GMM (Blaise and Dillenseger, 2023) and two used propensity score matching (Bom & Stockel, 2021; Stöckel & Bom, 2022). One used targeted maximum likelihood estimation (Torres et al, 2021), one used DID (Methi et al, 2024), one used generalised estimating equations (Zhao et al, 2023) , one used adjusted linear mixed models (Kirvalidze et al, 2023), and one was unclear (Angelini & Costa-Font, 2023).

I was particularly interested in the three studies that included both the caregiving effect and the family effect because they build on the work of Bobinac (Bobinac et al, 2010; 2011) and are more able to capture the full impact of caring on HRQoL, are so are likely to be informative for my analysis. Each of these is discussed in turn.

3.4.4.5 Bom et al (2019)

Bom et al used data from four waves of panel data from the Dutch Study on Transitions in Employment, Ability and Motivation (STREAM) to analyse the mental and physical health impacts of providing informal care, and to try to disentangle the family and caregiving effects. The authors selected a subsample of individuals who may potentially provide informal care to a parent or partner, including 4,400 males and 3,528 females at baseline and with 17,055 male and 13,693 female observations across all waves. To account for factors that might determine whether somebody becomes a caregiver, the authors used a first-difference model to control for time-invariant personal characteristics, condition on lagged health status of the caregiver, and include remaining time-varying factors as variables in the model. They estimated the resulting dynamic panel data model using GMM.

The change in caregiver's health at each year is dependent on the change in their health status observed in the previous year, and on the change in the informal care provision status at time t, and on the change in the health state of a spouse or close family member, and on the vector of changes in individual time-varying characteristics.

Caregiver health was measured using four self-reported outcomes (in separate models): the SF-12 PCS and the MCS, the SF-36 vitality subscale to measure fatigue, and the Center for Epidemiological Studies-Depression-10 (CES-D) scale to measure depression.

The authors found a statistically significant negative caregiving effect for the SF-12 MCS and a statistically significant negative family effect for the SF-12 MCS and depression. When the family effect was excluded, the size of the caregiver effect for SF-12 MCS increased. When stratified by caregiving intensity, the authors found a larger impact on SF-12 MCS and vitality for a greater intensity of caregiving. They did not find a significant effect for SF-12 PCS. The caregiving effect was largest when caring for a spouse, and this particularly affected vitality and depression. There was also a negative family effect when a spouse became ill, which was not found for other family members. The authors reported that for all subgroup analyses, the effects (except for SF-12 PCS) were larger for females than males and often significantly different. The authors compared their GMM model with results from OLS and compared with models that allow for fixed effects, and found the estimates of caregiving and family effects were not substantially different.

The authors concluded that their research highlighted the importance of separating the family and caregiver effects, and of considering subgroups given the significantly different effects for different subgroups.

3.4.4.6 Han et al (2021)

Han et al used seven waves of biennial data from the US Health and Retirement Study to understand the relationships between spousal activity limitations (activities of daily living (ADL) and instrumental activities of daily living (IADL)), caregiving behaviour and depressive symptoms. They used a sample of 6,475 couples (dyads), with 57,844 person-wave observations.

They used three longitudinal actor-partner interdependence models, each of which allowed for asymmetrical effects transitioning into and out of caregiving, and each decomposed into within-person (fixed effects) and between-person (random effects) components. Their preferred model considered the effect of transitions into and out of caregiving and their interaction with type of caregiving on CES-D depressive symptoms, adjusted for spousal activity limitations. They found both wives and husbands experienced fewer depressive symptoms when they transitioned into caregiving than when they remained non-caregivers, in the context of their spouses' activity limitations. In a simpler model that excluded spousal activity limitations, they found that transitioning into caregiving increased depressive symptoms in wives, but not in husbands. When they considered spousal activity limitations and transitions into caregiving, but not the interaction between them, they found that spousal activity limitations increased depressive symptoms in wives but not husbands, and that transitioning into caregiving did not significantly affect depressive symptoms for husbands or wives. The authors suggested that the findings of the simpler models were confounded by omitting the interaction between the type of spousal limitation and transitions into caregiving, since spousal health problems lead to both the need for spousal caregiving and depression among spouses.

The authors concluded that their research demonstrated that caregiving itself does not increase depressive symptoms, and caregiving may actually alleviate the emotional health consequences of having a spouse afflicted with a disability.

3.4.4.7 Han (2023)

Han used nine waves of biennial data from the US Health and Retirement Study to understand the relationships between maternal disability and dementia and caregiving behaviour and depressive symptoms in adult children. They used a sample of 4,812 people whose mothers were alive prior to the observation period and included them until the wave after their mother died, leading to 18,442 person-wave observations.

They used multilevel models with asymmetrical fixed effects for transitions into and out of caregiving (the effect of transitioning into caregiving is not assumed to be identical in size and opposite in direction to the effect of transitioning out of caregiving) and estimated a series of within-between random effects models. Within-between random effects models consider the “between effects” of key differences between population groups and the “within effects” of variability within individuals over time. In their preferred model which included maternal health, characteristics of the mother and family and interaction between these and caregiving transitions, the authors found that transitioning into caregiving alleviated the negative effects of moderate maternal disability (effects for severe disability showed a similar trend but were not statistically significant, and there was no such relationship for dementia). Other models which excluded maternal health conditions or additional characteristics found that transitioning into caregiving was associated with increased depressive symptoms. The authors concluded that their research demonstrated that transitioning into caregiving did not increase depressive symptoms once the effects of maternal health conditions are considered, and that caregiving can alleviate the detrimental effects of moderate maternal disability.

Table 1: Summary of longitudinal studies

Author (year) country	**Data source**	**Study design / methods**	**Exposure(s) and measurement**	**Health outcomes**	**Main findings**
Angelini & Costa-Font (2023) Europe	Survey of Health, Retirement, and Ageing in Europe	Short and long run, outcomes as a function of partner cancer diagnosis	Caregiver: helped spouse or partner (who was diagnosed with cancer).	EURO-D,self-rated health, pain, health behaviours	Short-term: cancer diagnosis has a significant effect on partner's self-reported health, effect is bigger for non-caregivers than caregivers.
Bertogg et al (2025) England	English Longitudinal Study of Ageing	Asymmetric fixed-effects	Caregiving: looking after anyone in the past week	Allostatic load using 9 biomarkers, cognitive function	Men: transition into caregiving improves allostatic load and memory. Women: small improvement in memory, but no stat sig with allostatic load. Women's memory function increases when caregiving

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
					intensity increases, men's no change Both: transition out of caregiving no change in outcomes.
Blaise and Dillen Seger (2023) Netherlands	Dutch Longitudinal Internet Studies for Social Sciences panel	GMM, two-step system using first lag of dependent variable	Caregiving: providing any care of care in the last 12 months	Life satisfaction	Caregiving significantly negatively affects life satisfaction
Bom et al (2019) Netherlands	Study on Transitions in Employment, Ability and Motivation	GMM dynamic panel model, first differences to control for time-invariant characteristics, condition on lagged health status	Informal caregiving: spent part of time in past 12 months on giving informal care. Spouse or family member become severely ill last year.	SF-12 MCS, SF-12 PCS, SF-36 vitality, CES-D	Statistically significant negative caregiver effect for the SF-12 MCS, and a statistically significant negative family effect for the SF-12 MCS and depression. When the family effect was excluded, the size of the caregiver effect

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
					for SF-12 MCS increased. They did not find a significant effect for SF-12 PCS
Bom & Stöckel (2021) UK and Netherlands	UK Household Longitudinal Survey, Study on Transitions in Employment, Ability and Motivation	Propensity score matching and regression Comparison of UK and Netherlands	UK: Informal caregiving: looking after someone living with you or providing regular help to someone not living with you. Netherlands: Informal caregiving: spent part of time in past 12 months on giving informal care. Spouse or family member become severely ill last year.	SF-12 MCS, SF-12 PCS.	Dutch caregivers experienced a statistically significant negative effect on MCS, whereas the effect for UK caregivers was insignificant. The effect for PCS was non-significant for Dutch caregivers, but statistically significant and positive for UK caregivers

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
Cheng et al (2023) China	China Family Panel Studies	Fixed effects ordered logit model	Caregiving: taking care of mother/father's household chores or meals	Life satisfaction/subjective wellbeing	Strong evidence of negative impact of informal caregiving on subjective wellbeing, bigger impact for high frequency care.
Costa-Font & Vilaplana-Prieto (2024) Europe	Survey of Health, Retirement, and Ageing in Europe	Instrumental variables (IVs) fixed effects filtered model	Co-resident caregiving: regularly helps someone in household with personal care. Non-residential caregiving: helps family member, friend or neighbour. IVs: being only child, eldest child, number of brothers, youngest child, number of sisters	Life satisfaction	Being a caregiver reduces life satisfaction

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
Eibich (2023) England	English Longitudinal Study of Ageing	Instrumental variables (IVs)	Caregiving: looked after anyone in the past week. IVs: presence of one living sibling, only one parent alive	Self-reported health, CES-D, binary ADLs and IADLs.	For men, informal care provision improves health (self-reported, ADLs and IADLs) but not depressive symptoms. Effects less for women.
Han (2023) US	Health and Retirement Study	multilevel models: caregiving transitions using asymmetrical fixed effects modelling approach, within-between random effects model	Caregiving, defined as spending a total of >=100 hours in the last 2 years helping their mother with basic personal activities. Maternal disability, defined by whether the mother needed help with personal needs and could be left alone.	CES-D	In models excluding maternal disability and dementia, transitioning into caregiving significantly increased depressive symptoms. In models including maternal disability and dementia, transitioning into caregiving was not

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
			Maternal dementia, defined by whether a doctor had ever said the mother had a memory-related disease.		associated with increased depressive symptoms. In models including interaction terms, transitioning into caregiving significantly alleviated the negative effect of maternal disability.
Han et al (2021) US	Health and Retirement Study	Actor-partner interdependence models. Asymmetrical fixed effects within-person and between-person effects	Spouse's activity limitations (ADL and IADL). Caregiving, where spouse had activity limitations and reported receiving help from spouse.	CES-D	Spousal activity limitations significantly increase depressive symptoms. Caregiving does not increase depressive symptoms when controlling for activity limitations.

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
					In models including interaction terms, caregiving significantly alleviated the negative effect of spousal activity limitations.
Henry & Cullinan (2024) Ireland	Growing up in Ireland	System GMM, outcome function of outcome at previous time points, health status of child and time-varying characteristics. GMM estimator using lagged first differences as "internal" instruments.	Presence and diagnosis of child's chronic physical or mental health problem/illness.	CES-D	Maternal health outcomes differ by the child's health condition, but even the largest effect is only significant at the 10% level.
Jain & Ma (2024)	English Longitudinal	GMM, linear dynamic	Spouse:	Respondent:	Significant cross-

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
England	Study of Aging	models. Single equation health production functions, one for each spouse, where vectors of health outcomes at time t are functions of own lagged health outcome vectors and spouse	Mobility difficulty index, CES-D, Cognitive health	Mobility difficulty index, CES-D, Cognitive health	domain and cross-spouse spillover effects which vary by health domain and gender. Wives' mental health significantly influenced by husband's mental health, husbands' mental health more influenced by wives' physical health.
Krivalidze et al (2023) Sweden	Swedish National study on Aging and Care in Kungsholmen	Adjusted linear mixed models	Caregiving: providing care to a friend or relative. Dual role: simultaneous received (formal or informal) care and provided informal care.	Limitations to life & perceived burden HAT objective health status	Having a dual role and caring for a spouse at baseline were associated with a faster average annual decline of the HAT score

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
					over the 12-year period
Krämer & Bleidorn (2024) Netherlands, Germany, Australia	Dutch Longitudinal Internet Studies for Social Sciences (LISS) panel German Socio-Economic Panle (SOEP) Household, income and Labour Dynamics in Australia Survey (HILDA)	Fixed effects	Caregiving: providing informal care	Wellbeing (life satisfaction, affect, depression/anxiety, loneliness)	Consistent wellbeing losses in women after becoming caregivers, and wellbeing losses for men that partly didn't extend beyond first year. More time caring led to worse wellbeing.
Labbas & Stanfors (2023) Europe	Survey of Health, Retirement, and Ageing in Europe	Fixed effects	Caregiving: providing care a parent	EURO-D depression and quality of life (CASP-12)	Caregiving relates to higher risk of depression among women in the Nordic countries.
Le & Ibuka (2023) Japan	Japanese Study of Aging and Retirement	Instrumental variables (IVs) fixed	Caregiving: current or anticipated	CES-D, life satisfaction, self-rated	Informal caregiving significantly

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
		effects	involvement in the care of your mother or father-in-law, and provision of help to family members, relatives, friends, neighbours. IV: mother- or father-in-law requires nursing care.	health, hand-grip strength, joint disorder	increased the probability of poor mental health and being unsatisfied with life, but did not affect physical health.
Methi et al (2024) Norway	Norwegian Counties Public Health Survey	DID	Caregiving: providing regular unpaid help or supervision to someone in need of help	Wellbeing: positive hedonic wellbeing, negative hedonic wellbeing, eudaimonic wellbeing, social wellbeing	New caregivers experience a decline in psychosocial wellbeing (positive and negative hedonic), but the decline begins prior to the caregiving transition non-significant but trend for contributing

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
					more to happiness of others
Stockel & Bom (2022) UK	UK Household Longitudinal Survey	Propensity score matching and regression. Dynamic matching procedure.	Informal caregiving: looking after someone living with you or providing regular help to someone not living with you. Caregiving intensity: hours per week.	SF-12 MCS, SF-12 PCS,	In the static matching procedure, statistically significant small immediate negative effects on MCS, which persisted up to 5 years, and a statistically significant small positive effect on PCS at the first time point, but no lasting effect. Dynamic model, providing informal care tends to have a negative impact on mental health.

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
Torres et al (2020) Mexico	Mexican Health and Aging Study	Pooled cross-sectional analysis, with lagged variables, using targeted maximum likelihood estimation	Caregiving: providing care to spouse with IADL	CES-D, lower body functional limitation, chronic health conditions	People providing care experience more depressive symptoms
Van den Berg et al (2014) Australia	Household, Income and Labour Dynamics in Australia	Fixed-effects ordered logit model	Weekly informal care time	Self-rated life satisfaction	Statistically significant negative effect of informal caregiving on life satisfaction, and a statistically significant effect of intensive caregiving.
Zhao et al (2023) China	China Health and Retirement Longitudinal Survey	Generalised estimating equations	Caregiving: spouse reported participant was primary caregiver	CES-D	People who transitioned into ADL caregiving had higher degree of depressive symptoms, people who transitioned into

Author (year) country	Data source	Study design / methods	Exposure(s) and measurement	Health outcomes	Main findings
					IADL did not differ from non-caregivers

ADL: activities of daily living, CASP:Control, Autonomy, Self-realization, and Pleasure ,CES-D: Center for Epidemiological Studies-Depression, DID: Difference-in-Differences, GMM: generalised method of moments, IADL: instrumental activities of daily living., SF-12: short-form 12, SF-36: short-form 36.

3.5 Conclusion

There is a stark contrast between the evidence base for carers' HRQoL and broader evidence on changes in carers' health or wellbeing. In recent years, the number of studies providing information on the HRQoL of carers for patients with specific health conditions has expanded rapidly, but the design of the studies has hardly evolved: studies are typically cross-sectional and non-comparative, so while they provide information on utilities for carers, they do not provide estimates of a causal effect of caring on HRQoL and they do not provide any information on carers' HRQoL trajectories. The handful of longitudinal studies that have been conducted have been over very short time periods (2 years or less) with few timepoints, and analyses have been limited in considering only the family or the caregiving effects. By contrast, there is much more longitudinal evidence on the relationship between transitions into and out of caregiving, patient health conditions, and carers' health or wellbeing outcomes. These studies have made use of observational datasets rather than conducting new studies and have included multiple waves of data. They have used a wide range of methods to try to isolate causal effects, including multilevel modelling, matching, and instrumental variables.

Taken in its entirety, this review suggests that caring does affect carers' health or wellbeing outcomes, but that the relationship is complex and there is a need to consider measures of patient health as well as the caregiving role. Much of the evidence has focussed on the relationship between caring and mental health (or wellbeing, or depression) and the evidence for the relationship between caring and physical health is much weaker and less clear. It may therefore be important to consider the relationship between caring and individual domains of HRQoL measures rather than the index score.

The paucity of longitudinal evidence on carers' HRQoL highlights the need for new analyses to identify the effect of caring on HRQoL. The much richer longitudinal evidence on carers' health or wellbeing outcomes demonstrates that it is possible to use existing observational datasets for this analysis (where relevant outcomes are included), showcases potential methods, and highlights the need for including information on patients' health and measures of caregiving. The next chapter of this thesis combines these two findings and analyses existing observational datasets to identify a causal effect of caring on HRQoL.

4 Estimating the effect of caring on health-related quality of life

4.1 Introduction

Chapter 3 highlighted the limited evidence on the causal effect of caring on HRQoL, and the scarcity of published longitudinal analyses of carer utilities. This chapter aims to address that evidence gap and provide estimates of carer HRQoL that can be used in economic models. Two separate analyses are summarised below.

The first analysis uses longitudinal data to examine how carers' HRQoL (measured using SF-6D) evolves over time. It draws on the concepts of the family effect and the caregiving effect, discussed in 2.6.1Section 2.6, using a household survey to link co-resident care-recipients and carers. The findings are published (Pennington et al, 2025b).

The second analysis studies the causal effect of different volumes of care by applying causal inference methods to a cross-sectional study of EQ-5D-3L and 5L. At the time of writing the thesis, a UK value set for EQ-5D-5L was not publicly available, so the analysis applies the mapping from 5L response to the 3L value set (Hernandez-Alava & Pudney, 2022).

Both of these analyses provide estimates of carers' HRQoL that can be used in economic evaluation. They draw on two different datasets, methods, and HRQoL measures so the estimates differ, and each has their own strengths and limitations. In this chapter, each analysis is described and discussed in turn, before a conclusion compares and discusses both analyses and their implications together. For the first analysis, I organised a public involvement workshop with carers to discuss the analysis and how it compared to carers' experiences, to ensure I had included all relevant variables and to aid interpretation of my findings. This is described briefly in the section about this analysis and in detail in the Appendix in Chapter 8.

4.2 Longitudinal analysis of carers' HRQoL in the UK Household Longitudinal Survey (UKHLS)

This section includes an overview of the analysis including aims and summary level findings (for more information please see the full text paper), information on the mapping from SF-12 to EQ-5D (not included in the full text paper), description of the public involvement work (mentioned briefly in the full text paper) and a discussion of the analysis.

4.2.1 Overview

I aimed to address the evidence gap in how carers' HRQoL changes over time through longitudinal analysis of UKHLS in "Unpaid Caring and Health-Related Quality of Life: Longitudinal Analysis of Understanding Society (the UK Household Longitudinal Survey)"

(Pennington et al, 2025b). To my knowledge, this is the first study to analyse carers' HRQoL, as measured by utilities for use in economic evaluation, at more than two time points.

I made use of two important features of UKHLS: firstly, its longitudinal design with 13 waves of data, and secondly, that as a household study it captures carer and cared-for dyads. This meant that I was able to predict changes in carers' HRQoL as a function of both the caregiving effect (measured by intensity and duration of caregiving) and the family effect (measured by the HRQoL of the cared-for) identified by Bobinac and colleagues (Bobinac et al, 2010; 2011). This is unique for a longitudinal analysis to date and is suitable for use in economic evaluations of interventions for patients where the patients' HRQoL is already captured, and caregiving duration and intensity can be measured from other sources.

I used fixed effects panel data analysis to estimate the effect of caring on HRQoL. Fixed and random effects models differ in their assumption about the correlation between the unobserved individual effect and the covariates: uncorrelated in the random effects model case and correlated in a fixed effects model. To establish which model to estimate, I used the Hausman test. The test rejected the null hypothesis of no systematic differences between the fixed- and random-effects estimators and, therefore, the fixed effects model was preferred in this case.

The main survey in UKHLS contains the self-completion SF-12 module. This consists of 12 questions that ask participants about: their general health, how their health limits activities, how their physical health affects work, how their mental health affects work/achievements, pain, feelings, energy and social life. These questions each report ordinal answers on a scale 1-5 (although sometimes 5 is best, and in other cases 5 is worst). To measure HRQoL, I used SF-6D, calculated from the SF-12 collected in UKHLS (Brazier and Roberts, 2004).

SF-6D is a generic preference-based measure of health and can be used within cost-utility analysis (see Section 2.4). The SF-12 domains used in calculating SF-6D are:

- Mental health meant accomplished less
- Felt down-hearted and depressed
- Health interfered with social life
- Physical health limited kind of work
- Pain interfered with work
- Health limited moderate activities
- Had a lot of energy

The SF-6D analysis found that carers' HRQoL decreased by 0.045 for every additional year spent caring, and that carers' HRQoL decreased by 0.0123 if the HRQoL of the cared-for decreased by 0.1. Caring at low, medium or high intensity had a statistically significant

positive impact on (increased) HRQoL, with similar effect sizes for all intensities. Duration of care had a statistically significant negative impact on (decreased) HRQoL and since this increased with each additional year of care, it negated the HRQoL gain from caring intensity within 2 years.

Different measures of HRQoL can produce different utility values, and NICE states a preference for EQ-5D (Dawoud et al, 2022). EQ-5D is not routinely collected in UKHLS, but was included (3L and 5L, see Section 2.4) for a subsample of UKHLS respondents in the Innovation Panel (a sample of 1,500 households where new survey methods and contents are tested) in wave 11 in 2019 (Hernandez-Alava & Pudney, 2022). It was therefore possible to develop a mapping from the SF-12 MCS and PCS scores included in UKHLS to EQ-5D, so that I could consider EQ-5D as a measure of HRQoL for both the carer and the cared-for, which could be used directly in economic models (described in the Appendix in Chapter 07).

The EQ-5D results (not reported within the published paper) were consistent with the SF-6D analysis: caring decreased HRQoL by 0.038 each year, and a decrease in the HRQoL of the cared-for of 0.1 corresponded to a decrease of 0.0084 in the carers' HRQoL. Coefficients for care intensity were also similar, positive and small.

In a public involvement workshop, carers discussed that caring could affect all aspects of HRQoL, with negative impacts on mental health and feelings of loneliness, as well as some effects on physical health. Some carers suggested there could be positive impacts of caring. This was valuable in understanding that the coefficients for caring could initially be positive, but that longer term effects were negative, particularly as some carers expressed that having a definite end to the caring role could be helpful. Further details of the public involvement workshop are described in the Appendix in Chapter 8.

4.2.2 Discussion

In my analysis, summarised above, I found that carers' HRQoL was correlated with patients' HRQoL and that carers' HRQoL declined as duration of caring increased (Pennington et al, 2025b).

My finding that carers' HRQoL continued to decline over time was consistent with the experiences of my public involvement group, but it is a new contribution to the literature, where models had previously focussed on intensity of care. This can be captured in economic models and would mean that interventions which reduce the duration of or delay the start of unpaid caring offer an additional HRQoL benefit to carers. However, I noted that it may not be appropriate to extrapolate this result over an extended period as relatively few carers in UKHLS provided care for more than five consecutive years. In economic models of

interventions for older people (such as Alzheimer's disease), a duration of five years caring may be appropriate, but carers of children with life-limiting conditions (for example Duchenne Muscular Dystrophy) may provide care for much longer than five years. I further explored this by considering lagged or initial HRQoL scores in regression analyses, which were statistically significant. This may suggest that future research should consider the use of dynamic models to predict changes in carers' HRQoL.

My conceptualisation of carer's HRQoL as a function of the duration and intensity of unpaid care ("caring for") and the HRQoL of the person they care for ("caring about") was consistent with previous (cross-sectional) analysis of carer's HRQoL (Bobinac et al, 2010; 2011), but depicts a unidirectional flow in which the care-recipient's HRQoL can influence the carer's HRQoL, but not the other way round. In the publication itself I noted this limitation, discussing that if carers' HRQoL can also influence care-recipient's HRQoL, then the analysis is subject to simultaneity bias and so likely represents an upper bound of the family effect ("caring for"). From a more theoretical standpoint, this "one-way flow" in care relationships has been criticized by both disability and care researchers as indicating a false dichotomy, and does not allow for the potential reciprocity in care relationships where the "care-recipient" can give as well as receive help (Kröger, 2009). The use of models which consider a bidirectional relationship between carer and cared-for HRQoL, such as actor-partner interdependence models (Han et al, 2021) or multilevel models with asymmetrical fixed-effects (Han, 2023), should be considered for future research. However, to include these in economic evaluation would also require that decision models give more consideration to the complexity of the relationship between carer and cared-for.

4.3 Causal inference methods to estimate the effect of caring on EQ-5D using cross-sectional data

4.3.1 Overview

I aimed to estimate the effect of caring on EQ-5D in a dataset that collected EQ-5D directly, since NICE states a preference for EQ-5D to measure carer's HRQoL (National Institute for Health and Care Excellence, 2022). This made use of an existing dataset that collected EQ-5D for a UK population and included questions about whether the respondent provided unpaid care to someone inside or outside their household, and how many hours per week (Hernández Alava et al, 2022)

This was a cross-sectional study: participants were only surveyed once. This meant I did not have data on how carers' HRQoL changed as their caring responsibilities changed, and I could not use the panel data analysis methods that I used in analysing UKHLS (Pennington et al, 2025b). This therefore required other methods to adjust for selection bias (people who

become carers may differ from people who do not become carers and these differences may also affect HRQoL), and so I used inverse probability weighting (IPW) to adjust for potential confounders to estimate the average treatment effect on the treated (ATT) where the treatment was providing unpaid care at differing hours per week. I adjusted for all the potential confounders for which data was available: age, sex, family status, education, and diagnosis of a health condition. The survey did not contain information on the care-recipient, so it was not possible to disentangle the family and caregiving effects, as I did in analysing UKHLS (Pennington et al, 2025b).

The survey asked participants to complete both 3L and 5L. At the time of writing, the UK value set for 5L was not publicly available, but a mapping was available to map 5L responses to 3L scores (Hernandez-Alava & Pudney, 2022). I analysed the effect of caring on 3L, and on 5L mapped to 3L. When the 5L value set becomes available, I intend to analyse the effect of caring on 5L and compare the ATT for 3L and 5L.

The ATT for 3L was statistically significant and negative for all weekly hours of care above 0-4 hours per week. The ATT for 5L mapped to 3L was statistically significant and negative for all weekly hours of care above 0-4 hours per week except for the “don’t know” category. The effect sizes for 3L were approximately -0.09 for caring 5-9 hours per week, -0.12 for caring 10-49 hours per week and -0.19 for caring 50+ hours per week. The effect sizes for 5L mapped to 3L were smaller: approximately -0.06 for caring 5-9 hours per week, -0.10 for caring 10-49 hours per week and -0.14 for caring 50+ hours per week. The difference in effect sizes between 3L and 5L mapped to 3L was partly explained by differences in the probability of participants reporting full health. Non-carers were more likely to report full health than carers using 3L and 5L, but non-carers were more likely to report full health using 3L than 5L. The relative difference between the proportion of carers and non-carers reporting full health was therefore bigger using 3L than 5L.

In a scenario analysis considering only people who provided unpaid care within their own household, the ATT increased in size and was statistically significant across all weekly hours including 0-4. Effect sizes broadly increased as the volume of caring increased, up to -0.25 for 50+ hours for 3L and -0.18 for 5L mapped to 3L.

4.3.2 Discussion

I found that unpaid caring above 4 hours per week was associated with a significant negative effect on EQ-5D-3L and 5L mapped to 3L. This negative effect was larger for higher weekly volumes of care, most notably 50+ hours per week. My finding is largely consistent with the literature suggesting that caring negatively affects HRQoL (Wittenberg et al, 2019), caring

has a greater impact above 50 hours per week (Carers UK., 2023), and EQ-5D disutilties from caring are around -0.1 (Theakston & Mott, 2025).

While I have shown that carers have lower EQ-5D than non-carers, the analysis is still subject to selection bias and does not prove that caring leads to worse HRQoL. I adjusted for all potential confounders that were recorded in the dataset, but there may be other unobserved differences between carers and non-carers, that I could not adjust for. My analysis also assumes that the IPW model is correctly specified. I considered that doubly robust methods which would also have used outcome regression (OR) and only required either the OR or IPW to be correctly specified were not appropriate for modelling EQ-5D since we know that EQ-5D scores are bounded and its distribution has characteristics that are not easily accommodated with a linear regression. A further limitation is that there is no information provided on the person being cared for – this means that the caregiving and family effects are combined into a single estimate, and it is unclear how generalisable this evidence is to other populations. However, it adds to the body of evidence suggesting caring negatively affects HRQoL and provides estimates that can be used in economic models.

4.4 Conclusion

The two analyses reported in this chapter indicate that there is generally a detrimental effect of caring on HRQoL, as measured by SF-6D and EQ-5D. Despite drawing on different data sources, applying different methods and using different HRQoL measures, they both reach the same conclusion: even after controlling for potential confounders, carers have statistically significantly worse HRQoL than non-carers. The effect of caring is multi-dimensional, affecting multiple domains of HRQoL.

The two analyses have different strengths and limitations. One analyses a rich source of longitudinal data on patient-carer dyads and so reports separately the family and caregiving effects but relies on a mapping to convert SF-12 to SF-6D. The other uses EQ-5D and a larger sample (of carers and non-carers) but does not report information about the care-recipient and is limited in its cross-sectional design. I anticipate that both could be useful for informing economic evaluation, depending on the priorities of the decision-maker.

Both analyses are for UK populations: while these may be ideally suited for use in economic evaluations for NICE, they may have limited transferability to other countries. This will be particularly the case where formal care provision and cultural expectations differ.

Now that I have established that caring does indeed have a detrimental effect on HRQoL, the next step is to consider exactly how this can be captured in economic evaluation. This will involve examination of existing models, and consideration of the family and caregiving

effects captured in the analysis of patient-carer dyads demonstrated in my analysis of UKHLS (Pennington et al, 2025b).

5 Methods for modelling patient and carers' HRQoL

5.1 Introduction

In Chapter 4, I demonstrated that unpaid/informal caring has a causal impact on HRQoL. Where interventions impact carers' HRQoL as well as patient's HRQoL, it is therefore important to model the change in carers' HRQoL to ensure that all relevant outcomes are included, so that population health can be maximised. Chapter 4 further provided results that can be used in modelling carers' HRQoL. This chapter now focusses on how economic evaluation can include carers' HRQoL, discussing the current methods and proposing a new approach.

As discussed in Section 2.7, the two most commonly used approaches are the "(absolute) utility" approach which includes carer's QALY until the patient dies, and the "disutility" approach which includes a QALY loss for carers until the patient dies. The "(absolute) utility" approach may be interpreted as implicitly assuming that patient death has an entirely negative impact on carers' HRQoL since the carer utility then becomes 0 and equivalent to death. The "disutility" approach may be interpreted as implicitly assuming that patient death has a wholly positive impact on carers' HRQoL since the carer disutility becomes 0, implying that carer's utility returns immediately to that of the general population. (In reality, these interpretations are somewhat nuanced, since there is also a case to be made for considering whether the impacts on carers are considered "relevant" after the patient dies, and in terms of the value placed on active or bereaved carers). One proposed solution to the carer QALY trap (whereby interventions that lead to extended patient survival lead to carer QALY losses, as per the "disutility" approach) is to also include a "bereavement effect" QALY loss for carers, after the patient dies (Landfeldt & Sandhu, 2024). If the bereavement effect is the same size regardless of when the patient dies, it becomes smaller for life-extending treatments due to discounting. Alternatively, the magnitude of the bereavement effect may vary depending on when the patient dies – in particular it may be larger where patients die as children than as adults.

This chapter is structured into four sections to help understand current and alternative modelling methods, and to consider the implications of including carers' HRQoL in economic evaluations.

1. To understand whether either the "(absolute) utility" or "disutility" method is appropriate, this chapter first explores theories from care research in the context of economic evaluation. This considers how care affects both the carer and the care-recipient and how this aligns with the current methods for including carers' HRQoL in

economic evaluation. This section concludes that methods for modelling carers' HRQoL must recognise the complexities of caring relationships.

2. The second section considers empirical evidence on how bereavement affects carers' HRQoL, to understand if there is any evidence to support either the implication that carers' HRQoL becomes 0 or is severely detrimentally impacted after the patient dies (as per the "(absolute) utility" approach), or that carers' HRQoL immediately returns to pre-caring levels after the patient dies (as per the "disutility" approach). This section concludes that there is little evidence to support a sustained bereavement effect for carers and concludes that incorporating a bereavement effect will not solve the challenges in modelling carers' HRQoL.
3. In the third section, I propose and describe a new approach for modelling carers' HRQoL, drawing on my research into the family and caregiving effects (in Section 4.2). This approach aims to better capture the complexity of caring relationships, and to separate out the effects of bereavement on the family effect (which may be negative, due to grief when someone the carer cares about dies, as per the "(absolute) utility" approach) and the caregiving effect (which may be positive, due to removal of the HRQoL loss due to the burden of caregiving, as per the "disutility" approach). This section discusses the theory and applies the approach to three case studies.
4. The fourth section considers whether explicitly modelling carers' HRQoL in some economic evaluations for specific interventions has implications for decision-making since displaced interventions may also have had carers' HRQoL impacts which were not explicitly considered. This includes a discussion of the arguments for and against considering displaced carers' HRQoL from the literature, and how and whether this should be considered in the context of my new approach proposed in the third section.

Taken in its entirety, this chapter provides a comprehensive overview of the rationale, methods and implications of including carers' HRQoL in economic evaluation.

5.2 Why is it so challenging to include carers' HRQoL in economic evaluation?

5.2.1 Current approaches to modelling carers' HRQoL

The current approaches for modelling carers' HRQoL, described in Section 2.7, include the "disutility" approach (including a negative impact on carers' HRQoL as long as the patient is alive), "(absolute) utility" approach (including carer utilities as long as the patient is alive), "incremental approach" (including an incremental carer utility/disutility relative to a baseline

health state) and “multipliers” (estimating carer QALY gains as function of patient QALY gains). All approaches make assumptions about how carers’ HRQoL changes in response to patients’ HRQoL changes and following the patient’s death, and each has limitations. The challenges in including carers’ HRQoL are not merely technical or mathematical but are at least partly due to fundamental differences in research into economics and research into care. In this section, I first discuss the differences underpinning the “ethic of care” and the “ethic of justice” to understand why it is difficult to incorporate care into economics. Next, I discuss the importance of considering “interdependence” in the context of the relationship between patients and carers. Then, I discuss why care has a negative effect on the carer (as per the “disutility” approach and identified in the evidence in Chapter 3 and Chapter 4). Finally, I consider whether we need to combine patient and carers’ HRQoL.

5.2.2 Ethic of care

Economic evaluation and care theory have different roots. Economic evaluation is rooted in the “ethic of justice” and the idea that resources should be distributed fairly such that everyone can be treated equitably, which requires consideration of opportunity cost and cost-effectiveness (Charlton, 2022; Cookson, 2015). This is rooted in the idea of individualism, that everyone is a utility-maximiser. In the specific context of health economics, this assumes that everyone affected by HTA decision-makers is motivated solely to maximise their own QALY gain.

By contrast, the “ethic of care” acknowledges that when people make decisions about themselves and other people, there may be conflicting outcomes (Gilligan, 1993). The stages (rectangles) and transition periods (thought bubbles) depict the decision-making process in Figure 5.

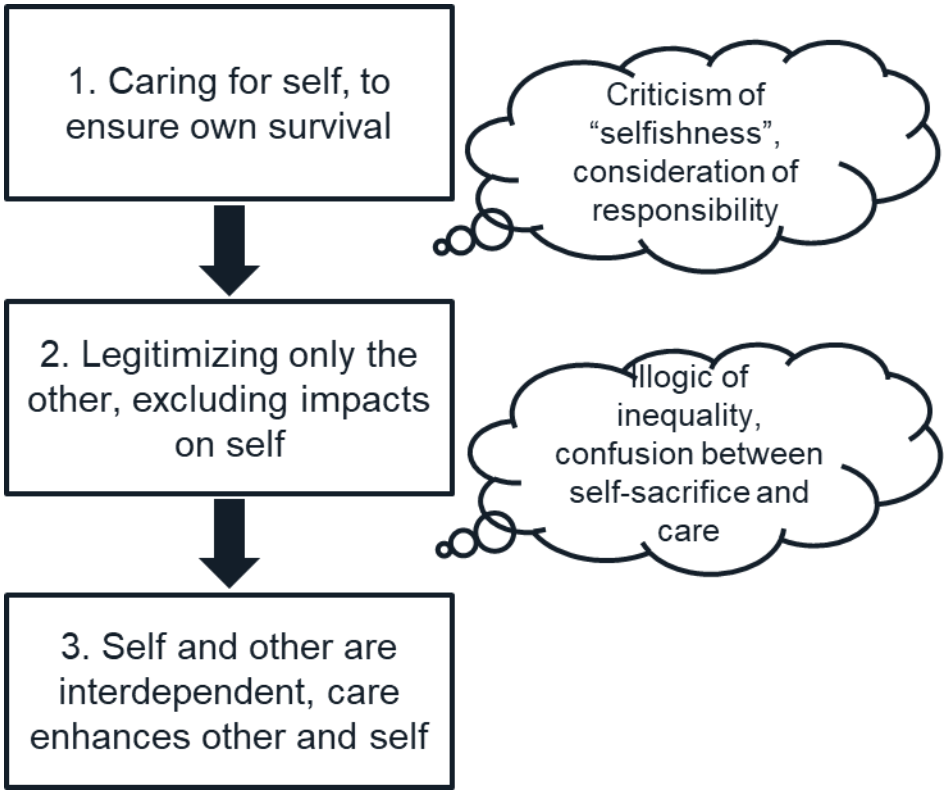


Figure 5: Ethic of care

When we consider the current methods of modelling carer's HRQoL in the context of the ethic of care, and with the assumption that providing care negatively affects the carers' HRQoL, the "disutility approach" where carers' HRQoL improves after the patient dies appears to fit within the first stage. Carers deciding about a life-extending treatment for the person they care for would be assumed to prefer the patient did not receive the treatment since the prolonged period of care is detrimental to their own QALYs. Stage 2 could be seen to represent either entirely excluding carer HRQoL from economic evaluation (only including impacts on the patient), or the "(absolute) utility" approach where carers' QALYs are only legitimate (included in analysis) as long as the patient is alive. The third stage, whereby carer and care-recipient are interdependent such that caring enhances the carer and care-recipient, has yet to be established.

5.2.3 Independence and interdependence

Previous research has discussed the challenges of reconciling disability studies (which focus on independence and the needs of disabled) with care research (which emphasises the interdependence between people) (Kröger, 2009). Patient-focussed health economics can be considered analogous to disability studies, and the importance of independence within this stream of research is emphasised by the questions in the commonly used HRQoL measures: EQ-5D asks respondents about problems in washing and dressing, usual

activities, and mobility, and assumes that fewer problems and increased independence is always preferable to requiring support. SF-6D similarly includes multiple domains relating to functioning and what the person can "do". Even capability focussed measures like the ICE-CAP A include a question about independence (University of Bristol., 2025).

Research into unpaid carers within health economics has focussed on the effect that caring has on the carer, and not on the care-recipient. This unidirectional flow has been discussed in Section 4.2.2. However, it is also unclear to what extent measures of HRQoL would identify the effect of caring on the care-recipient, for example the questions in EQ-5D do not specify whether the respondent has problems in the specific domains without assistance, or when considering support from informal carers.

Methods for including carers' HRQoL in economic evaluation have so far assumed that the effect of caring is either entirely negative (as in the "disutility" approach) or entirely positive (as in the "(absolute) utility" approach and excluding carers beyond patient death). Consideration of the interdependence between patients and carers may include recognition that the relationship between patient and carer HRQoL is complex and can have positive and negative components.

5.2.4 Does care benefit the carer?

The third phase of the ethic of care whereby care benefits the carers is at odds with the evidence in Chapter 3 and Chapter 4 where caring negatively affects HRQoL. If this is the case, then perhaps this should be interpreted as a critique of the societies that have made it impossible for unpaid carers to maximise their own utility. In her essays on caring for her disabled daughter, philosopher Eva Feder Kittay describes a concept of "doulia", in which care is provided for carers, since advocating for better care for patients must necessarily require better care for carers too (Kittay, 2001). The challenges unpaid carers face are well documented (Carers UK., 2024) and better support to address these could include financial support, support to stay in employment, provision of respite care or better access to support for mental and physical health.

5.2.5 Do we need to combine patient and carer HRQoL?

Given the different roots and concepts underlying economics and care, it is not surprising that combining them has caused challenges. Rummery notes that:

"Attempts to codify care and bring it within the public domain of 'justice' run the risk of divorcing it from its ethical and emotional roots" (Rummery & Fine, 2012)

One alternative to attempting to model the interdependence of patient and carers is Frye's "loving eye" (Frye, 1983). Frye posed this as an alternative to the "arrogant eye" in which an

oppressor sees the oppressed only in relation to themselves. The relative roles of patient and carer are oppressor and oppressed are not important here, the important concept is that the two parties are separate. Under the “loving eye”,

- The carer’s own interests are not denied.
- The carer and patient are separate.
- The carer and patient have different interests which are not blended.
- The patient and carer are independent of each other.
- The carer does not try to fit the patient’s interests into their own world view or simplify them.

This approach may support reporting patient and carer QALYs separately and not aggregating them. While this is tempting because it does not require us to make any trade-offs between patients and carers, it does not help decision-makers maximise outcomes for society.

5.2.6 Conclusion

Caring is a complex subject, in which carers cannot always be assumed to be either utility-maximisers or entirely altruistic. The relationship between carers, the person they care for and their HRQoL is therefore also complex and its application in economic evaluation warrants careful consideration. Methods to include carer HRQoL in economic modelling must recognise the complexities of these relationships.

5.3 Effect of bereavement on carers’ HRQoL

5.3.1 Overview

This analysis aimed to address the evidence gap of how carers’ HRQL changes when the person they cared for died. This is important in understanding whether the assumptions that carer HRQoL returns to full health (as per the “disutility” approach) or that carer HRQoL becomes equivalent to death (as in the “(absolute) utility” approach) are reasonable. This is also an important concept in understanding the appropriate time horizon for modelling carers’ health outcomes, as highlighted in the Spillovers in Health Economic Evaluation and Research (SHEER)Taskforce Recommendation 4:

“Family and caregiver health spillovers should be measured over a time horizon sufficient to capture all relevant impacts of the health condition(s)/intervention(s), including the full impact of caregiving duties, bereavement, or both…” (Henry et al, 2024)

Inclusion of a HRQoL detriment following bereavement has also been suggested as a solution to the carer QALY trap (Landfeldt & Sandhu, 2024), and has been considered in

some NICE appraisals (National Institute for Health and Care Excellence, 2018b; National Institute for Health and Care Excellence, 2021).

My analysis made use of the longitudinal household study design in Understanding Society, which naturally captures the HRQoL of bereaved household members, and also records where co-residents were caring for each other (Pennington et al, 2025a). One of the arguments against including the effect of bereavement on HRQoL in economic evaluation is that people are bereaved when loved ones die in all circumstances, regardless of whether they were caring for the deceased. I therefore wanted to consider the HRQoL effects of bereavement on carers, and on people whose family member died but they were not caring for them.

I conducted separate analyses for carers and non-carers, and in each case used DID to compare people who became bereaved to people who did not become bereaved. It was important to compare against people who did not become bereaved since HRQoL may change for other reasons, such as ageing.

I wanted to be able to include all bereavements (and comparisons) across the 12 waves of data that were available (Wave 13 could not be analysed for this purpose since it could not report the variable of whether the person was bereaved by the next wave), with bereavement events happening each year. However, it was important to be able to make comparisons between people who were bereaved and people who were not bereaved within any given wave, since there are other time-specific factors that could influence people's HRQoL trajectories (for example covid, recessions, changes in government). I also wanted to consider multiple time periods to study HRQoL trajectories following bereavement (rather than only in the first year after bereavement). To account for these factors, I used Callaway and Sant'Anna's method for DID with multiple time periods and variation in treatment timing (Callaway & Sant'Anna, 2021).

I used the doubly robust approach, combining OR and IPW to adjust for potential confounders: age, sex, educational qualification, household income and presence of long-standing illness. I assumed conditional parallel trends, and while I could not formally test for these, analysis of the HRQoL trajectories pre-bereavement suggested they were reasonable, and there was no reason to believe people could select into bereavement. I also assumed limited treatment anticipation – while this is a standard assumption in DID, it was debatable in the context of "anticipatory grief", which was suggested to apply to carers more than non-carers (Schulz, 2008). In my analysis, I found a very small but statistically significant SF-6D loss for bereaved carers compared to non-bereaved carers in the year before bereavement (-0.00762), whereas this was not significant for non-carers. Comparison

of SF-6D changes for co-resident non-carers suggested that the HRQoL loss prior to bereavement for carers was at least partially due to anticipatory grief, as well as an increased caregiving burden.

I found a statistically significant HRQoL loss in the year immediately following bereavement for both bereaved carers (compared to non-bereaved carers or non-bereaved non-carers) and bereaved non-carers (compared to non-bereaved non-carers). I did not formally compare the two groups – while methods for Difference-in-Difference-in-Differences exist (DDD), their application in this context is not straightforward, particularly given the differential treatment timing and multiple time periods. However, comparison of the two DIDs suggested that the short-term effect of bereavement was similar for carers (-0.0283 (95% CI − 0.0432 to − 0.0134)) and non-carers (-0.0383 (95% CI − 0.0527 to − 0.0239)). In the long-term, the effects were not statistically significant, suggesting no lasting impact of bereavement on HRQoL. Furthermore, the effect sizes were small, so even if they were applied in economic evaluation for bereaved carers, the impact on the overall QALYs and cost-effectiveness would be small.

5.3.2 Discussion

I concluded that there was no evidence of a differential effect of bereavement on HRQoL for carers and non-carers, and that the small short-term bereavement effect would have little impact if it were included on cost-effectiveness analysis. In my paper I proposed that, contrary to suggestions by Landfeldt and colleagues (Landfeldt & Sandhu, 2024), including bereavement effects was not the solution to the carer QALY trap.

In the discussion section of my paper (Pennington et al, 2025a) I compared my findings to published literature reporting the HRQoL of bereaved individuals/couples, noting that while some did find a significant impact compared to non-bereaved individuals, they were limited in their cross-sectional design and comparisons with matched controls or the general population. A cross-sectional study of parents whose child died of Spinal Muscular Atrophy (SMA) found that the average SF-6D for people whose child died <5 years ago was 0.64, and those whose child died 5-10 years ago was 0.70, and those whose child died >10 years ago was 0.73 (Lavelle et al, 2025). These values were lower than age- and sex-matched population norms of 0.86, 0.84, and 0.82. While the comparison with population norms is challenging as different measures were used (HUI for population norms) and there may be other underlying differences, the data do suggest that bereaved parent's HRQoL is lower when the child died more recently. This may support a longer duration of the bereavement effect than my study found. These differences may arise because of the nature of the caring relationship and the condition: the impact of the death of a child may be very different to the

death of a spouse/parent. Theoretically, if interventions were able to extend the life of a child such that they did not die in childhood, this bereavement effect may be avoided. But if interventions only postponed death by a few years, the only reduction in the bereavement effect would be due to discounting. Furthermore, bereavement effects to other family members much further into the future are not well researched – for example if the child reached adulthood, had a child and then died, there may be a bereavement effect to the child of losing a parent. While I recognise that bereavement may have a greater impact on HRQoL in some scenarios than identified in my analysis, it seems unlikely that the inclusion of a bereavement effect on HRQoL in economic evaluation is the solution to the carer QALY trap. I was therefore compelled to explore further how economic evaluation can realistically capture both impact of caring on HRQoL, and the impact of life-extending treatments on carers.

5.4 A new approach for modelling carers' HRQoL

5.4.1 Overview

I describe a new approach for including carers' HRQoL in economic evaluation (Pennington et al, 2026). Previous research (see Section 2.6.1) has analysed carers' HRQoL as a function of the HRQoL of the person they care for (the family effect) and the amount of care they provide (the caregiving effect), and my own analysis in Section 4.2 estimate these effects in longitudinal analysis. My modelling approach uses the concepts and numerical estimates of the family and caregiving effects to model carers' HRQoL in economic evaluation. The family effect is positive: as patients' HRQoL increases, carers' HRQoL increases. The caregiving effect is negative: as the amount of care that carers provide increases, carers' HRQoL decreases. Modelling the two effects separately allows a trade-off between the benefits of extending patient survival (applying the positive family effect over a longer time) against the negative HRQoL impact of increasing the caregiving burden (applying the negative caregiving effect over a longer time). Traditional approaches have assumed that the effect of caring is either wholly positive (as in the "(absolute) utility" approach) or wholly negative (as in the "disutility" approach).

The family effect has parallels with the "(absolute) utility" approach, in assuming that carers experience HRQoL gains when either the HRQoL of the person they care for improves, or when the person they care for lives longer. An important difference is that the family effect considers the *effect* on HRQoL whereas the "(absolute) utility" approach considers the carers' entire HRQoL, including aspects unrelated to caring. The QALYs arising from the family effect are therefore expected to be much smaller than those arising from the "(absolute) utility" approach.

The caregiving effect has parallels with the “disutility” approach, in assuming that carers experience HRQoL losses when their caregiving duties increase and that they experience gains in HRQoL when they no longer have to provide care. Whereas disutilities capture any differences between carers’ HRQoL and the general population due to either the family effect or the caregiving effect or other factors unrelated to caring, the caregiving effect isolates the specific HRQoL loss due to the burden of caregiving. The QALYs, or indeed QALY losses in the case of extended patient survival, from the caregiving effect are therefore expected to be much smaller than those arising from the disutility approach. It is hypothetically possible that carers actually experience a gain in HRQoL as a result of providing care, once the family effect is controlled for – two studies in the review in Chapter 3 supported this (Han et al., 2021, Han, 2023). In this case, carers would experience a HRQoL gain from providing care. However, most evidence suggests a negative HRQoL effect of providing care.

Interventions that do not affect patient mortality but improve patient morbidity lead to a QALY gain for carers through all approaches since:

- Carer utilities are higher so there are more carer QALYs in the “(absolute) utility” approach
- Carer disutilities are smaller so there are fewer QALY losses in the “disutility” approach
- Patient HRQoL is higher so the family effect leads to gains in carer HRQoL in the family effect
- Carer burden is decreased so the caregiving effects leads to smaller carer HRQoL losses in the caregiving effect.

Interventions that extend patient survival have different implications for carer QALYs in the four approaches since:

- Carer QALYs are included for longer so there are more carer QALYs in the “(absolute) utility” approach
- Carer QALY losses are included for longer so there are fewer carer QALYs in the “disutility” approach
- Patient HRQoL is positive and non-negative for longer leading to gains in carer HRQoL in the family effect
- Carer burden lasts longer leading to more carer HRQoL losses in the caregiving approach.

The current dichotomy of the “(absolute) utility” approach versus the “disutility” approach has meant there has been no possibility of trading off the carer QALY gains from increasing

patient survival with the carer QALY losses from extending the duration of caregiving. Combining the caregiving effect and the family effect allows us to make that trade-off, and to make transparent how the carer HRQoL spillovers have arisen.

In my publication, I apply my new approach, and the "(absolute) utility" and "disutility" approaches to three case studies, to explore the feasibility and impact of each approach. I chose three case studies of economic evaluations where carers' HRQoL has often been considered relevant, and where sufficient data was available to perform the calculations for each approach. The data I required was the time that the modelled patients spent in each health state, the carers' utility values, the patients' utility values, and a way to estimate the family and caregiving effects. I chose two NICE appraisals in rare chronic paediatric conditions that have caused controversy due to the carer QALY trap: spinal muscular atrophy (SMA) (National Institute for Health and Care Excellence., 2023b) and Duchenne Muscular Dystrophy (DMD)(National Institute for Health and Care Excellence., 2023a). The documentation for NICE appraisals is thorough and detailed, and in these cases the data was unredacted. I chose another case study in a different health condition at the other end of the life course: Alzheimer's Disease. This is an area where there is a wealth of data on carers' HRQoL, and an open-source model was available which provided most of the required data (Handels et al, 2024). All of the case studies provided the time patients spent in each health state, carers' and patients' utility values.

Ideally, the family effect and caregiving effect would be estimated from a study of patient-carer dyads as in my analysis (Pennington et al, 2025b). The family effect would be the coefficient for patient HRQoL (or change in patient HRQoL) and the caregiving effect would relate to a measure of caregiving burden as appropriate for the patient health condition and captured within the economic model. However, in the three case studies within the paper, I demonstrated how the family effect can be transferred from other data sources and combined with carer (dis)utilities to estimate the caregiving effect. The limitation here is the assumption that the family effect is generalisable across populations, caregiving relationships, and health conditions. Further research is needed to determine whether this is true.

Consideration of the family and caregiving effect is also useful for highlighting potential issues or assumptions within the carer utility data. For example, in the DMD case study, patient utilities were treatment dependent (lower for comparator than intervention) but carer utilities were treatment independent. This could only be true if either the size of the family effect differed between treatments or if the caregiving burden was higher for the intervention than the comparator. Since this does not seem plausible, it highlights that a treatment that is

anticipated to affect patient HRQoL within a given health state should also be expected to affect carer HRQoL. This would support using treatment-dependent carer utilities where treatment-dependent patient utilities are used. In the SMA case study, the reported carer disutility was smaller than the disutility that would be expected from the family effect assumed from other studies. This could happen if the family effect had been overestimated or if there was a positive effect of caregiving in this population, but neither seem plausible given the nature of SMA and the precedents from other health conditions. The more likely explanation is that the combination of the patient and carer utility data was not valid – they came from two different studies in different populations and used different measures of utility. Decomposing carer HRQoL into the family and caregiving effects can therefore highlight where there are issues with the underlying data. These issues would apply equally across all modelling approaches but are revealed in the family and caregiving effect approach.

At a public involvement workshop, unpaid carers and representatives of caring organisations recognised the limitations with the current methods for modelling carers' HRQoL and were broadly supportive of my new approach. However, they felt that the term "family effect" was misleading as it implied this was an effect on all family members beyond the primary carer, but we were unable to agree on a better term. They also expressed some concern regarding the generalisability of the size of the emotional impact, particularly in transferring the effect size between different populations and relationships (for example parent carers of children versus spousal carers of older adults) and between chronic and acute conditions. I was encouraged that my approach offered something new and is potentially valuable, but that it needs further research across conditions and interventions. A full report of the public involvement workshop is provided in the Appendix in Chapter 10.

5.4.2 Discussion

In my three case studies, the incremental carer HRQoL effects estimated using the family and caregiving effect were always positive (intervention leads to a HRQoL benefit for carers) but always much smaller than either the incremental patient QALYs or the carer QALYs estimated using the (absolute) utility approach.

However, there may be scenarios in which the negative caregiving effect is larger than the family effect, and in this case the carer HRQoL effects from extending patient survival may be negative. This may represent the "worst case scenario" discussed by Kröger (Kröger, 2009), and it may be true that in this scenario, the carers' HRQoL would genuinely be improved by allowing the patient to die. However, this finding will also arise where the total caregiver disutility is large, and the family effect has been underestimated. It is therefore likely to be valuable to conduct further research to estimate the size of the family effect in

specific populations, for example where parents are caring for children with life-limiting conditions (such as DMD and SMA) and/or where carers are caring for people with high levels of disability. The carer disutility in these scenarios is likely to be somewhat higher than in my analysis of patient-carer dyads (Pennington et al, 2025b), or in the literature considering family HRQoL related to acute childhood conditions (Al-Janabi et al, 2016c). Consideration should also be given to the scenario where the carer utility is lower than the patient utility – for example where a young child is first diagnosed with a life-limiting condition in the early or mild stages of the disease. In this example, the child's development may be on par with peers who do not have the condition, and they may be unaware of any future health impacts. But the parent carer is conscious that the child's health will deteriorate, and already begins to feel the mental or emotional impact of this, even before they need to start providing care. In this scenario, if the family effect disutility is calculated as a proportion of the patient disutility, then the caregiving effect must be particularly large and negative, for the total carer disutility to be larger than the patient disutility. In a more severe health state, the patient disutility may be larger than the carer disutility – assuming the same family effect size as in the mild state where the carer disutility is larger than the patient disutility implies that the caregiving effect must be larger in the milder state. However, the actual burden of providing care in the state where the patient health state is mild is expected to be less than in the state where the patient health state is severe (the parent does not need to provide care until the child's condition begins to develop). Furthermore, it seems counter-intuitive that extending survival in the best patient health state would have a larger negative impact on carers' HRQoL (through the greater caregiving effect) than extending survival in worse health states. In this case there may be an argument for including additional effects on the caregiver, for example a one-off impact of diagnosis that is applied to intervention and comparator in the initial model cycle/time point.

Attendees at the public involvement workshop questioned whether the family effect would apply to the wider family and not just the primary carer. In principle, the family effect could apply to individuals who care about the patient but are not actively involved in caregiving. Mathematically this would involve multiplying the HRQoL effects resulting from the family effect by the number of family members. However, evidence for the family effect stems from dyads where carers were actively providing care to the patient at least at one time point, and so it is unclear whether the family effect can be applied to non-caregiving family members. However, there is evidence that spillovers extend beyond a single close family member but with decreasing size. Al Janabi et al found a spillover of 0.18 for the closest family member and 0.11 for the second closest family member, and proposed using arithmetic or geometric progression to aggregate the spillovers for all family members (resulting in a total spillover

effect for the entire family of 0.33 or 0.48) (Al-Janabi et al, 2016c). It may therefore be appropriate to reduce the size of the family effect for additional family members, but the magnitude of this is unclear without also collecting data on the second closest family member.

It is possible that at least to some extent, the family effect exists within all families or relationships, even where one person is not providing care for the other. Conceptually, this is because people care about others and their HRQoL will be influenced by that of people they care about, even in the absence of a caregiving relationship. Mathematically, the family effect, as measured in analyses like mine and those of Bobinac (Bobinac et al, 2010; 2011) measures the correlation between the HRQoL of two co-resident people. This may arise from either party caring about the other, but it is also feasible that both people's HRQoL changes as a result of factors that apply to both people – for example one household member loses their job, or both household members become bereaved, or a recession or pandemic happens. In these cases, we would identify a correlation between both people's HRQoL that was not caused by a change in the health or HRQoL of either party. If the family effect, as measured in our analyses, exists within non-caring relationships, then it should be included in all economic evaluation even where patients do not have unpaid carers. However, the family effect, as it has been identified in my research and that of Bobinac (Bobinac et al, 2010; 2011), is derived from a population where one person cares for the other at some point in time. There is therefore no evidence to support the existence of the family effect, or to assume its magnitude is the same, in relationships where neither party provides care for the other. Furthermore, it is estimated in models that also consider the caregiving effect, and so the two effects should be considered in combination and applied only to populations where both are relevant.

So far, this chapter has discussed methods for including carers' HRQoL and proposed a new approach to do so. The next section considers the implications of including carers' HRQoL in economic evaluation in specific cases, but where there may be carers' HRQoL benefits from other interventions that have not explicitly been considered, and may be displaced in favour of funding an intervention where the benefits for patient and carers have been formally captured.

5.5 Displaced carer HRQoL

5.5.1 Introduction

If we are to make decisions about healthcare interventions that affect unpaid carers and include those specific HRQoL effects to carers, then Al-Janabi, Brouwer and McCabe agree that we should at least in theory consider the HRQoL effects resulting from displaced

interventions (Al-Janabi et al, 2016b; Brouwer, 2019; McCabe, 2019). All three also agree that estimating displaced carer HRQoL effects is challenging, but then disagree about the implication of this, with McCabe arguing that this potential bias is a reason to exclude carer HRQoL effects from any economic evaluation, and Al-Janabi and Brouwer advocating for further research into estimating carer HRQoL effects across interventions to address this.

5.5.2 Methods to estimate displaced carer HRQoL

McCabe discusses that to fully understand the evidence base for incorporating carer HRQoL effects into economic evaluation requires an atlas on the carer HRQoL burden across the spectrum of disease and socioeconomic characteristics. He argues that this would allow analysts to determine how the carer HRQoL effects from the new intervention compared to the average (McCabe, 2019).

Brouwer argues that we do not know what gets displaced when we make decisions on patient health, and so a lack of knowledge on carer HRQoL effects should provide an incentive to learn more about them so that we can better estimate average displaced carer HRQoL effects, and that including them in economic evaluation is the best way to learn about them. Brouwer advocates for more research into carer HRQoL effects, including measurement, valuation and inclusion of this in economic evaluation (Brouwer, 2019).

Drawing on their research into multipliers to represent the ratio of carer health effects to patient health effects discussed in Section 2.7, Al-Janabi et al discuss that if the ratio of carer HRQoL effects to patient effects is constant across new and displaced interventions then decision-making based on patient health effects alone is sufficient. This would be analogous to the family effect existing with the same magnitude across all interventions and the caregiving effect being constant and zero (or directly in proportion to patient health and therefore already captured in the family effect). Where carer HRQoL effects are not constant across the health system (either the family effect, or the caregiving effect, or both varies across interventions) then incorporating carer HRQoL effects into the decision-making process is essential to maximise health benefits. In order to avoid bias towards new interventions which include carers' HRQoL effects, the HRQoL effects that would have arisen from displaced interventions must also be known (Al-Janabi et al, 2016b).

5.5.3 Displaced carer HRQoL in the context of my research

The paucity of evidence on how caring affects HRQoL, highlighted in Chapter 3 demonstrates that we are a long way from completing the "spillover atlas" of carer HRQoL effects across health conditions and sociodemographic characteristics. While there is evidence for carer HRQoL across conditions, for example the 80 studies identified by Wittenberg and colleagues in 2018 (Wittenberg et al, 2019), these do not cover all

interventions available in healthcare systems. Instead, they suggest that carer HRQoL has been more heavily studied in some health conditions such as Alzheimer's and other types of dementia than in other health conditions. This may be because caring is particularly prominent or has a bigger burden in some health conditions than others – this would align with the suggestion by Al-Janabi et al that health spillovers are not constant across conditions (Al-Janabi et al, 2016b).

My analysis of UKHLS patient-carer dyads did not differentiate between patient health conditions, since the way that this information is recorded in UKHLS means that it does not accurately capture current or ongoing health conditions (Pennington et al, 2025b). I considered a scenario analysis using the limited information available, which was only available for one time point. This meant that patients were defined as either having a condition or not having a condition, there was no within-person variability in developing a health condition. I therefore used a random effects model to consider the variability between groups, rather than the fixed effects model used in the base case analysis. This scenario suggested carers had statistically significantly lower SF-6D scores only where the patient was diagnosed with asthma, and higher SF-6D scores where the patient was diagnosed with chronic heart failure (other non-statistically significant health conditions were arthritis, coronary heart disease, angina, myocardial infarction, stroke, emphysema, bronchitis, hyperthyroid, hypothyroid, liver condition, cancer, diabetes, epilepsy, hypertension, multiple sclerosis and HIV). I did not consider that this was particularly convincing evidence that carers' HRQoL differed by patient health condition, but I also did not consider it particularly robust given the format of the questions and typically low number of people for each health condition. It did not consider the severity of the condition (although this may be related to hours of care) or whether the diagnosed condition was the reason that the carer was caring for the patient.

If the family effect is generalisable across health conditions (and the non-significance of specific health conditions is an artefact of the data), then it may be possible to use the family effect as a multiplier/ratio, as proposed by Al-Janabi et al (Al-Janabi et al, 2016b). Excluding the caregiving effect from the analysis, I found a family effect of 0.109, which would suggest a ratio of 1.109 for total patient and carer health effects relative to patient health effects only. If we assumed that this was applicable to all health conditions, and that the population studied in my analysis was generalisable, in that most people with a health condition will receive care from a household member at some point, then we could use this ratio to adjust (reduce) the decision-making threshold. If the threshold was £20,000/QALY, this would become approximately £18,000/QALY (£20,000 divided by 1.109) However, if this ratio was the same across all new interventions too, then as noted by Al-Janabi et al, the decision-

making process would be exactly the same as if carer HRQL effects had been excluded from both new and displaced interventions (since an intervention which had an ICER of £20,000/QALY including only patient health effects would have an ICER of £18,000/QALY when additionally including carer health effects, using the same ratio) (Al-Janabi et al, 2016b).

Including only the family effect, either for new or displaced interventions, assumes that carers' HRQoL can be entirely predicted by patient's HRQoL. Evidence suggests that this is not true: Al-Janabi et al identified six mechanisms that affect carer's wellbeing and patient health is only one of them (Al-Janabi et al, 2019) and my own analysis found the duration of care and volume of care were statistically significant in addition to the family effect (Pennington et al, 2025b).

It may be tempting to assume that the family effect applies across all new and displaced interventions, but that the caregiving effect is an additional consideration in some specific new interventions. In this case, we could either adjust the threshold by the ratio of total health effects relative to patient health effects (for example reducing it to £18,000/QALY) and include the family effect in all new ICER calculations, or we could equivalently leave the threshold unchanged and exclude the family effect from all new interventions. In either case, we could include the additional caregiving effect for specific interventions where it was considered relevant. However, if we did this, assuming the detrimental impact of caregiving on HRQoL, then we would effectively be using a smaller version of the disutility approach. The implication would be that life-extending treatments would negatively impact carers' HRQoL, and the challenges associated with this have been discussed in Section 2.7 and Section 5.2.1.

The importance of accounting for displaced carer HRQoL effects depends on the magnitude of these likely effects: clearly if they are expected to be large then the impact of not considering them will be greater. My case studies considered three conditions where carers' HRQoL was believed to be particularly relevant, but under my proposed approach, the carer QALY effects were less than 12% of the patient QALYs, suggesting that while the impact is not negligible, carer HRQoL contributes much less to decision-making than patient HRQoL (Pennington et al, 2026). It is also important to consider whether displaced interventions are likely to have spillover effects on carers. Carer HRQoL effects were only considered in <5% of NICE Technology Appraisals to date, suggesting that they are not commonly considered relevant (Kanters et al, 2024). Furthermore, they were typically focused on specific disease areas such as dementia and multiple sclerosis, which, to the best of my knowledge do not account for the majority of NHS expenditure on prescribing. The NHS programme budgeting

categories for 2023/24 indicate that the biggest area of spending is endocrine disorders (24.26% total) with diabetes contributing 13.95%, circulation disorders (16.42%), respiratory disorders (12.29%) and gastrointestinal disorders (11.5%) (NHS Business Services Authority., 2025). With the possible exception of stroke, these disease areas are not the same as disease areas where there is evidence for particularly large carer HRQoL effects (Wittenberg et al, 2019). Neurological disorders, musculoskeletal conditions and mental health conditions which are typically expected to have greater carer HRQoL effects are associated with much lower NHS spending (5.50%, 2.73% and 6.03% respectively). Considering all of this evidence together suggests that relatively few currently funded interventions would have an impact on carers' HRQoL, and the impact of including carers' HRQoL would be relatively small. Furthermore, HTA bodies such as NICE have not demonstrated any precedent for adjusting the threshold when evidence suggests it may not reflect the true value of displaced interventions (Claxton et al, 2015; Lomas et al, 2019; Naci et al, 2025), and while their manual permits committees to consider an increased threshold to account for additional factors such as severity, uncaptured benefits and equality, it has not stated that these same factors should be considered for displaced interventions (National Institute for Health and Care Excellence, 2022).

The inclusion of carer HRQoL effects is much more common within NICE's highly specialised technologies programme (67%) (Kanters et al, 2024), which may indicate that they may be expected as standard for these interventions, in which case should be considered for displaced interventions, but these use a different decision-making process and cost-effectiveness threshold so should be considered separately. If all interventions assessed by the HST process are expected to have carer HRQoL effects then the threshold should perhaps be adjusted to reflect this, but the origins of this threshold are unclear (National Institute for Health and Care Excellence, 2022).

5.6 Conclusion

It is important to include carer HRQoL effects in economic evaluation to reflect the outcomes of all affected individuals, but doing so is challenging. This is partly due to the disconnect in theories underlying economics and care, and partly due to the way that carer HRQoL has been conceptualised in economic evaluation to date. The "(absolute) utility" approach assumes that extending the duration of care positively affects carer QALYs, whereas the "disutility" approach assumes that caring can only ever negatively affect carer QALYs. My new proposed approach offers a trade-off between the HRQoL gains to carers from extending the life of the person they care for, and the negative HRQoL impact of increasing the caregiving burden. I propose using my new approach for new interventions and recognise that while displaced carer HRQoL effects are relevant, the impact is likely to be

small and we can learn more about displaced carer HRQoL effects by embedding methods for including carer HRQoL in new economic evaluations.

6 Discussion and Conclusions

6.1 Introduction

This thesis aimed to understand how economic evaluation currently includes carers' HRQoL, review the current evidence for the effect of caring on HRQoL, analyse the effect of caring on HRQoL, and develop methods for including the effect of caring on HRQoL in economic evaluation. Existing evidence suggested that carers have worse HRQoL than non-carers. My research was motivated by theory and an evidence base suggesting that patient health interventions had the potential to improve carers' HRQoL and aimed to address weaknesses that limited the application of existing evidence within economic evaluation.

My thesis included a background on the context for including cares' HRQoL in economic evaluation, exploration of care theory, a literature review, econometric analysis, review and development of modelling methods and consideration of implications. This chapter reflects on my research, outlines the relationships among the chapters, and discusses the implications for further research.

6.2 Key findings

The literature suggests that, on average, carers do have worse HRQoL than might be expected of the general population. Generally, there is evidence of a causal negative impact of caring on health and wellbeing outcomes, particularly relating to mental health, but much less evidence that caring causes a detrimental impact specifically on HRQoL. This is important when considering its application and potential impact in economic evaluation, where it will be assumed that changing the patient's HRQoL or amount of informal care required will directly impact carers' HRQoL.

Care theory, and the more sophisticated econometric models used to analyse longitudinal data on carers' health/wellbeing, demonstrate the complexity of the relationship between caring and health/wellbeing outcomes. Some studies found that while living with a person in ill-health lead to worse health/wellbeing outcomes, the act of providing care mitigated some of these negative effects. It is important to allow for both positive and negative relationships between care and wellbeing, and to consider separating the effect of providing care ("caregiving effect") from the effect of living with or caring about a person in ill health ("family effect").

My analysis of SF-6D from UKHLS, with its longitudinal design and wealth of information of patients and carers, found that the weekly volume of care was less impactful than the duration of care (years), and that there was also a statistically significant relationship between changes in patients' and carers' HRQoL (family effect). The use of longitudinal

analysis to identify a causal effect strengthens the evidence base and supports the conceptual arguments that there are causal negative impacts of both increased care duration and patients' HRQoL on carers' HRQoL. My analysis indicated that interventions that either reduce the delay the onset of unpaid care (or decreased its duration) or improve patients' HRQoL could also improve carers' HRQoL. This provides a rationale and data for including carers' HRQoL in economic evaluation with more flexibility to relate carers' HRQoL to changes in provision of care and patients' HRQoL.

My analysis of EQ-5D from a cross-sectional study confirms the findings of the longitudinal analysis, providing evidence that caring has a negative impact on NICE's preferred measure of HRQoL. This analysis found that there is a detrimental impact of caring on HRQoL, and that higher volumes of caring lead to statistically significantly worse HRQoL, even after adjusting for confounders. This analysis demonstrates that it is possible to identify a causal impact of caring from cross-sectional data and provides EQ-5D specific HRQoL decrements that can be applied in economic models to link the volume of unpaid care to carers' HRQoL.

The current methods for modelling carers' HRQoL in economic evaluations impose restrictions that the impact of caring on HRQoL must either be entirely positive (the "(absolute) utility" and "multiplier" approaches) or entirely negative (the "disutility" approach). These do not reflect the complexity of care, and neither are fully aligned with carers' experiences. There had been suggestions in the literature that inclusion of a "bereavement effect" could solve some of the problems related to modelling carers' HRQoL, if the carer disutility following bereavement were sufficiently large to mitigate increase in utility from no longer providing unpaid care. My analysis of the relationship between caring, bereavement and HRQoL suggested that bereavement has a small, short-term impact on HRQoL, and that the effect was similar for people who were or were not caring for the person who died. Inclusion of a bereavement effect was therefore unlikely to be a perfect solution of modelling carers' HRQoL.

Inspired by my findings in the longitudinal analysis that both the family and caregiving effects have a causal impact on HRQoL, I developed a new approach for modelling the impact of caring on HRQoL. My new approach considers the impact of caring on HRQoL via the family and caregiving effects and therefore allows a trade-off between the positive and negative impacts of caring. This is rooted in the concepts of care theory and can be included in economic evaluation using existing data and refined through further data collection. This approach does not impose restrictions on the direction of effect of caring on HRQoL, allowing the relationship between caring and HRQoL to vary across health states and over time.

6.3 Interpretation of findings

It is impossible to consider the combination of the care theory, existing evidence, my econometric analysis, and the different modelling methods, without concluding that the relationship between caring and HRQoL is complex. This is true not only in terms of identifying predictors of changes in carers' HRQoL, but also in terms of what is important to carers in their caring relationships.

While my econometric models find that there is a relationship between patients' HRQoL and carers' HRQoL, and between the volume or duration of care provided, these should not be interpreted as representing the totality of determinants of carers' HRQoL. Rather, they demonstrate that carers' HRQoL is influenced by many factors related to their caring relationships and responsibilities.

There is a real danger of oversimplifying the quantitative evidence and people's experiences and concluding that caring can only ever negatively affect HRQoL. It is essential to appreciate that caring relationships and experiences are multifaceted to be able to interpret the totality of the evidence base and to reflect this in economic evaluation. Reducing the effect of caring to a one-directional impact on HRQoL has caused problems in economic evaluation, which cannot be solved without fully appreciating the complexities inherent within caregiving relationships. My analysis of the relationship between bereavement and HRQoL refutes the suggestion that inclusion of a bereavement effect solves these problems.

6.4 My findings in the context of published literature

My research supports the view that caring impacts HRQoL and makes a difference to ICERs when included in economic evaluation. My estimates of the causal effect of caring on utility are consistent with published evidence, as is the size of the family effect identified in my longitudinal evidence.

My analyses add to the existing body of literature and provide evidence of a causal relationship between caring and HRQoL. The UKHLS analysis is unique in its longitudinal dyadic analysis and lends support to the idea that changing patient's HRQoL can change carers' HRQoL. The UKHLS analysis provides numerical estimates that can be applied in an economic evaluation where the patient HRQoL and duration of caregiving are known. The family effect estimate can be multiplied by the patients' HRQoL in any economic evaluation to calculate the family HRQoL effect on the carer, and the disutility that can be applied for each additional year of caring. These can be applied in any health condition where adult carers are providing care to adult patients. This offers flexibility beyond the currently available estimates from the literature which are specific to patient health states.

The analysis of EQ-5D provides estimates of the causal impact of caring at different volumes on HRQoL using NICE's preferred measure in a UK sample. These estimates have a higher level of granularity than is commonly seen in the literature (for example, carer versus non-carer) and the disutilities can be readily applied in economic models where the volume of unpaid care provided is known or can be estimated. This offers a source of carer HRQoL disutilities that conform to NICE's reference case and can be consistently applied across evaluations.

My proposed modelling method and associated data challenge the size of the difference that including carers' HRQoL makes to results and decisions about cost-effectiveness: I find a much smaller change in ICERs when carers' HRQoL is included using my approach. However, my proposed approach is consistent with theory that care is complex, and that decision-making is not straightforward: it does not imply that carers must choose between maximising either their own utility or that of the care-recipient. While there has been much discussion in the literature about the need for guidance on modelling methods, there has been little discussion on the relative merits of different approaches or how to design a new approach that overcomes existing challenges. By taking a step outside of HTA and looking at economic evaluation from the perspective of care theory, I have been able to develop an approach that provides new insight into modelling the relationship between caring and HRQoL.

6.5 Limitations

This thesis is subject to the limitations associated with the HRQoL measures, as well as the application of the cost-utility analysis framework in this context. I used EQ-5D where possible and another generic preference-based measure of health (SF-6D), consistent with NICE's stated preferences for measuring and valuing the HRQoL of carers (National Institute for Health and Care Excellence, 2022), but carer-specific instruments may be more sensitive to capturing HRQoL outcomes for carers (McLoughlin et al, 2020). Furthermore, there are many other effects of caring that are not captured when only considering HRQoL impacts to carers (for example the financial impact of caring) (Carers UK., 2024).

The review of the evidence aimed to be comprehensive whilst also taking a pragmatic approach but was not a fully systematic literature review and may therefore have missed some relevant papers. I consider it is unlikely that anything transformational was not identified, but there may be other studies that add to the body of evidence that have been omitted.

The analyses were both associated with limitations, primarily due to the availability of the data. Richer data on specific health conditions would have improved the UKHLS longitudinal

analysis and enabled consideration of the generalisability of the family and caregiving effects between patient disease areas. The cross-sectional EQ-5D analysis was limited by only studying the carer and not care-recipient, and at one point in time only. However, I used the best available methods to address these, including extensive sensitivity analyses.

The application of my new method for modelling carers' HRQoL in economic evaluation was limited to three case studies as it required information on patient and carer utilities and the time spent in each modelled health state (disaggregated life years). Disaggregated life years are not routinely reported in published economic evaluations and are often redacted in NICE submissions, which restricted the case studies I could access. They may therefore not necessarily be the optimal case studies but were chosen pragmatically and to represent a range of disease areas.

6.6 Future research

I have demonstrated that analyses of the relationship between caring and HRQoL need to consider patient HRQoL and different aspects of caregiving responsibilities and developed a method for using such data in economic evaluation. The challenge now is in generating this evidence across different conditions and contexts, or in determining transferability.

Ideally, every economic evaluation that includes carers would be informed by longitudinal analysis of carers' HRQoL, related to the patients' HRQoL and specific aspects of caregiving related to that intervention. This data would come from longitudinal dyadic studies which measure a range of variables and ideally also consider the HRQoL of family members who do not actively provide care. However, this is unlikely to be feasible for every economic evaluation, and so instead I suggest prioritising research that can be transferred between economic evaluations where data is currently sparse.

Specifically, future research should consider analysis of patient-carer dyads in conditions and relationships under-represented in my analyses, for example parent carers of children with chronic conditions. This will help determine whether the family effect is transferable between different conditions, whether there are additional drivers of caregiving effect in different scenarios, and whether the caregiving effect ever mitigates the family effect (has a positive effect) as suggested by some of the literature.

6.7 Conclusion

Caring statistically significantly affects HRQoL, but the relationship is multifaceted and complex. There are serious gaps and limitations within the evidence base to understand and accurately predict how patient health interventions will affect carers' HRQoL. However, research in this area is evolving and I demonstrate that it is possible to analyse existing

datasets to generate evidence for use in economic evaluation. There is a need for further data and analysis to address evidence gaps, particularly relating to the HRQoL of parent carers of children with long-term health conditions.

It is important to include carers' HRQoL effects in economic evaluation to reflect the outcomes of all affected individuals, but doing so is challenging. This is partly due to the disconnect in theories underlying economics and care, and partly due to the way that carers' HRQoL has been conceptualised in economic evaluation to date. The current commonly used approaches have serious limitations because they oversimplify the relationship between caring and HRQoL. My research demonstrates the scope to develop more flexible modelling methods that consider both positive and negative aspects of care. These methods now require more data, or evidence of transferability of data across contexts.

If the HTA community is serious about improving the HRQoL of unpaid carers, then we must spend time and resources in ensuring this is accurately measured and modelled. Advancing methodological developments simultaneously with collecting more data is essential for integrating this research into routine economic evaluation and maximising the HRQoL of patients and their carers.

7 Appendix: Mapping from SF-12 to EQ-5D

The Innovation Panel Wave 11 contains the EQ-5D-3L and the EQ-5D-5L for a sample of 1,622 people. I chose the EQ-5D-3L as our outcome variable of interest because the current 5L value set is not recommended by NICE (National Institute for Health and Care Excellence). The distribution of EQ-5D-3L scores is shown in Figure 6.

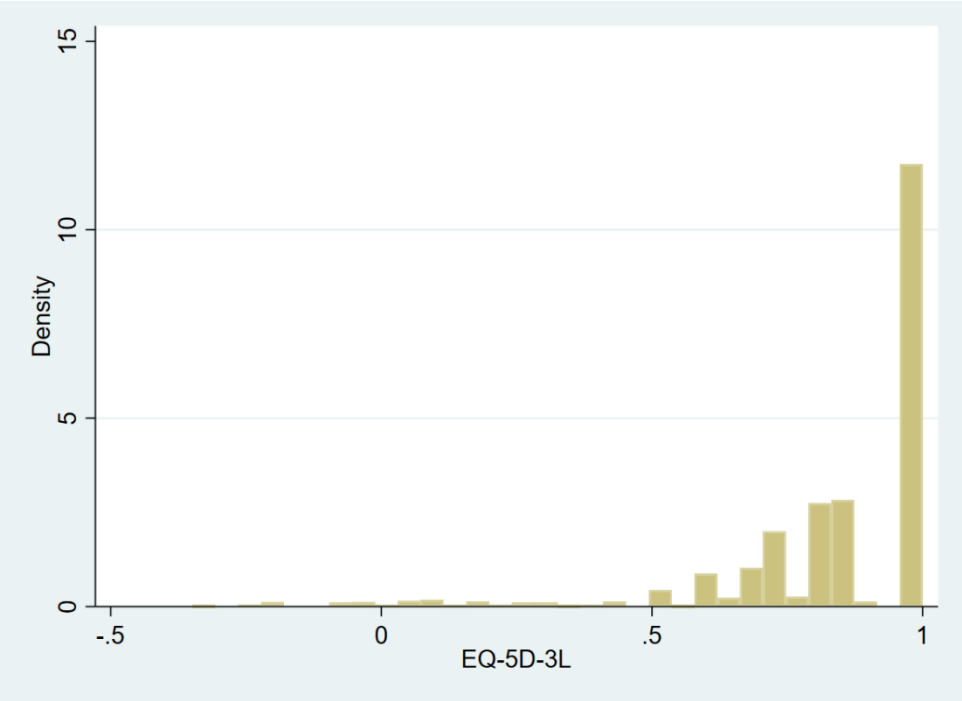


Figure 6: EQ-5D distribution in Innovation Panel

The SF-12 Mental Component Summary (MCS) and Physical Component Summary (PCS) scores are calculated from the answers to the original questions, resulting in continuous scales ranging from 0 (worst) to 100 (best). A slightly surprising quirk of the scales is that it is not possible to have very high scores in both the MCS and the PCS component. Respondents who give the best possible answers to all SF-12 questions will score PCS of 56.6 and MCS of 60.8. Reporting slightly worse health for some of the questions will lead to an increase in PCS and a decrease in MCS (for example, reporting feeling downhearted/blue some of the time) whereas reporting slightly worse health for other questions would lead to an increase in MCS and a decrease in PCS (for example, reporting pain interfering with work). This means that when we consider individually the relationship between either MCS or PCS and EQ-5D, we see what appears to be a surprising result, that the very highest scores of either MCS or PCS do not correspond to the highest EQ-5D scores. This is not a concern in my mapping that uses both MCS and PCS together to predict EQ-5D.

I used adjusted-limited dependent variable mixture models. My preferred model (based on AIC and BIC) has three classes/components, and class membership is determined by SF-12PCS and MCS scores. One of the classes is a "class of 1s" , which corresponds to full health and an EQ-5D score of 1. This reflects the distribution of the EQ-5D, and the relatively high proportion of people who report full health. The other classes have means of 0.83 and 0.35. An individual person's score in either of these classes is determined by their SF-12 MCS (linear and squared) and PCS scores (linear and squared), and age and sex. The model estimates are shown in Table 2.

Table 2: Estimates from mapping function

		b	standard error	z	P value	95% lower confidence interval	95% upper confidence interval
component 1: coefficients	age/10	-0.003	0.015	-0.190	0.850	-0.033	0.027
	male	0.014	0.032	0.440	0.661	-0.049	0.077
	sf12mcs/10	0.047	0.043	1.090	0.276	-0.037	0.131
	sf12mcs/10 squared	0.000	0.007	0.020	0.981	-0.013	0.013
	sf12pcs/10	0.086	0.054	1.590	0.112	-0.020	0.192
	sf12pcs/10 squared	0.002	0.007	0.230	0.815	-0.012	0.015
	cons	-0.349	0.144	-2.430	0.015	-0.631	-0.068
component 2: coefficients	age/10	-0.006	0.002	-2.440	0.015	-0.010	-0.001
	male	-0.001	0.006	-0.240	0.812	-0.013	0.010
	sf12mcs/10	0.030	0.021	1.440	0.149	-0.011	0.071
	sf12mcs/10 squared	0.000	0.002	0.140	0.890	-0.004	0.005
	sf12pcs/10	0.037	0.019	1.910	0.056	-0.001	0.075
	sf12pcs/10 squared	0.004	0.002	1.810	0.070	0.000	0.008
	cons	0.395	0.057	6.980	0.000	0.284	0.505
probability of being in component 1	sf12mcs/10	-2.561	0.320	-8.000	0.000	-3.189	-1.934
	sf12pcs/10	-2.653	0.305	-8.690	0.000	-3.251	-2.055
	cons	22.781	3.088	7.380	0.000	16.729	28.834
probability of being in component 2	sf12mcs/10	-1.342	0.278	-4.820	0.000	-1.887	-0.797
	sf12pcs/10	-1.568	0.265	-5.910	0.000	-2.088	-1.049
	cons	15.556	2.898	5.370	0.000	9.877	21.235
lns 1		-1.990	0.117	-17.070	0.000	-2.219	-1.762
lns2		-2.591	0.060	-43.090	0.000	-2.709	-2.473
sigma1		0.137	0.016			0.109	0.172
sigma 2		0.075	0.005			0.067	0.084

The cumulative distribution function is shown in Figure 7. The mean observed and predicted EQ-5D scores by SF-12 MCS and PCS are shown in Figure 8 and Figure 9. These figures

demonstrate that the predicted (or model) estimates are very similar to the actual (mean data), in terms of the distribution of EQ-5D scores and the relationship between SF-12 MCS/PCS and EQ-5D.

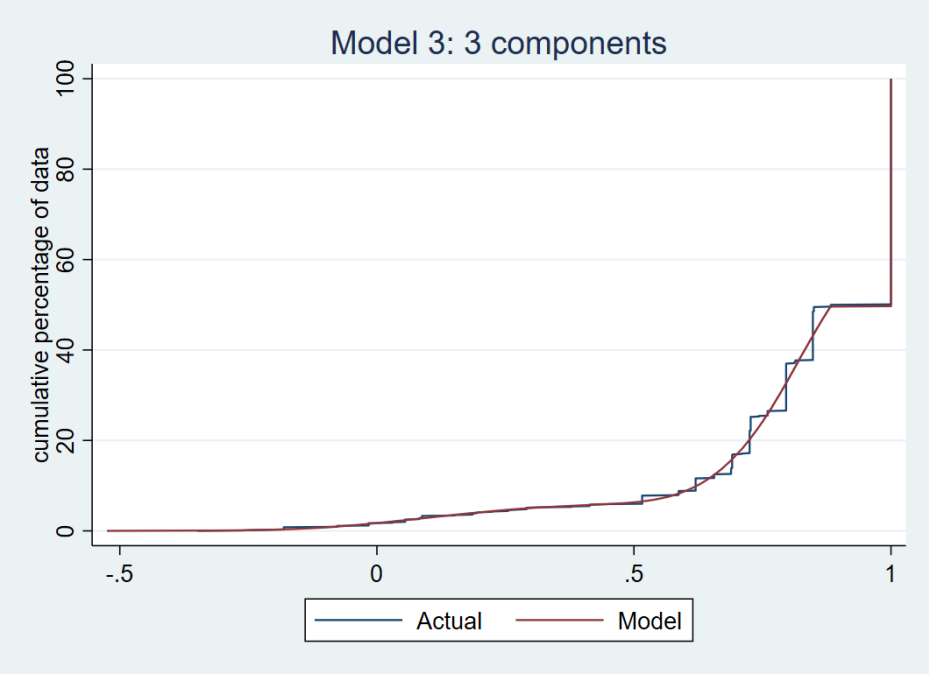


Figure 7: Cumulative Distribution Function: actual and modelled EQ-5D

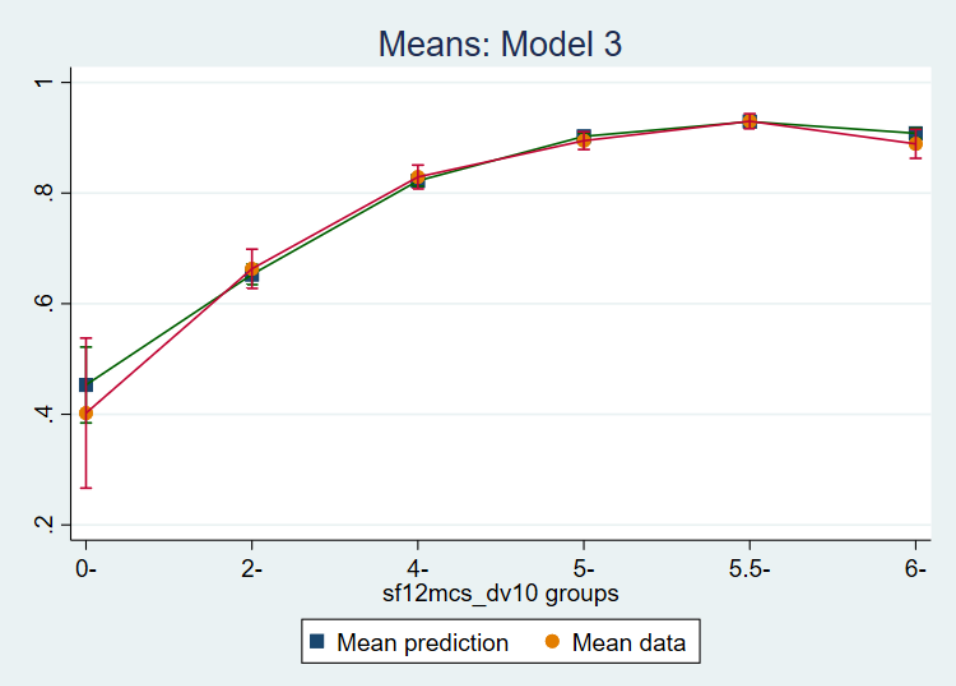


Figure 8: Mean observed and predicted EQ-5D by mental component summary score

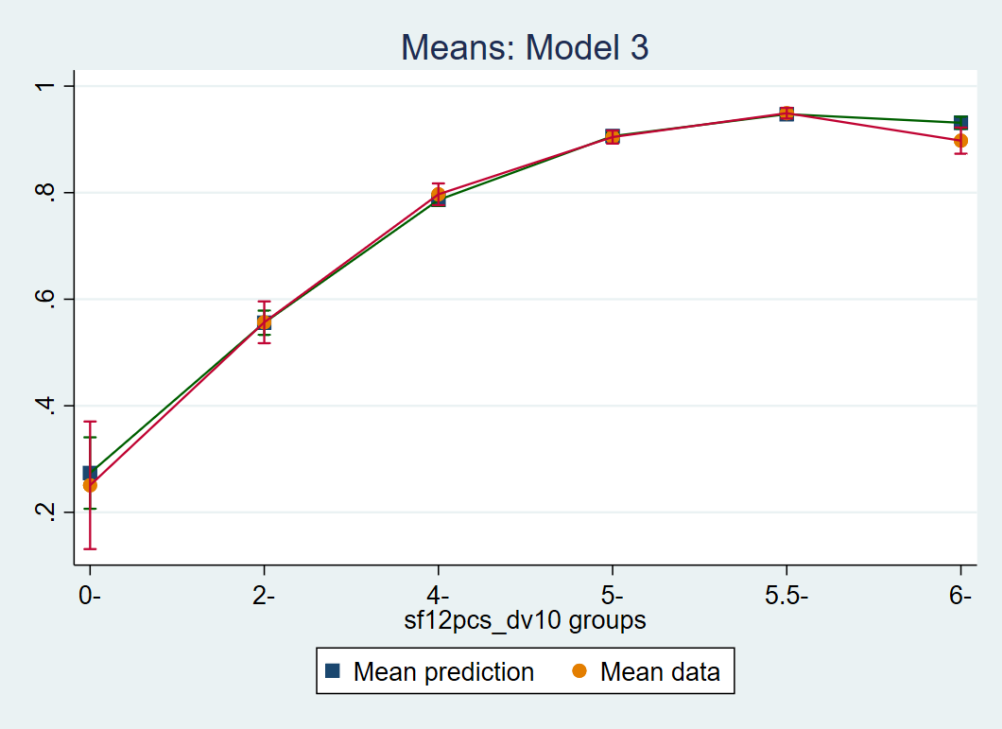


Figure 9: Mean observed and predicted EQ-5D by mental component summary score

8 Appendix: Public and patient involvement and engagement: UKHLS analysis

8.1 Introduction

In January 2024 I organised an online workshop with eight unpaid carers to understand their lived experience of caring and whether my analysis summarised in Section 4.2 and detailed in Pennington, Hernández Alava & Strong, 2025 (Pennington et al., 2025b) was consistent with this. This section provides a report of the summary of the main themes that were discussed at the public involvement workshop.

8.1.1 Set-up

The workshop was conducted using Google Meet. The workshop lasted 90 minutes. We used Google Docs to make notes (these were shared with participants during the meeting and anyone could contribute).

8.1.2 Participants

The public involvement workshop had twelve attendees: eight carers and four academic researchers.

The carers had a range of different experiences of caring in terms of the relationship to the person (or people) they care(d) for, the conditions of the person they care(d) for, and how long they were/are a carer for. Some carers lived with the person/people they cared for, others cared for someone living in their own home, others cared for people in assisted living or care home settings. Some of the people cared for had been in hospital for substantial periods of time.

8.2 Discussion: caring and health-related quality of life

I asked carers how being a carer affected health-related quality of life, specifically in relation to the domains of the SF-6D.

Participants noted that caring could affect every aspect of health-related quality of life, including all of the domains of SF-6D.

There was a strong feeling that caring can affect social functioning, with participants noting that caring responsibilities may restrict opportunities for having a social life, and many participants reporting feeling lonely or isolated by their caring role.

Many participants described an impact on their mental health and/or feeling downhearted or depressed. Some had been formally diagnosed with depression or anxiety. Others reported worrying or not being able to relax. Many participants referred to the “stress” of caring, but it is unclear how this fits into SF-6D and if it is fully captured.

Some participants described an impact of caring on physical health and pain. In some cases, this was through stress, anxiety or burnout which can manifest in physical symptoms (such as headaches). Some participants reported that they “cannot get ill” or that their own health conditions were not prioritised.

Carers also reported feelings of guilt and frustration.

8.3 Discussion: volume of care and potential positive impacts of caring

There were some discussions on what we meant by “care” and suggestions that it may be difficult to define or recognise when caring starts and what is or is not classed as “caring”. Some people described that they identify as carers constantly, even if they are not actively with the person. Many carers described constantly being responsible. There were discussions that the type of care might also be important, and that this may differ for different volumes of care. There was also discussion of formal care and the interaction or administration associated with arranging this.

Participants suggested that caring could have positive impacts, including:

- Spending time with someone that you love and the opportunity to get to know someone better
- Knowing that the needs of the person (whom the carer knows best) are being met
- Satisfaction in doing everything you could
- Growing in empathy, compassion, patience, resilience

Some carers described caring as a “privilege” and some would not describe themselves as carers, but by their relationship to the person they care(d) for.

8.4 Discussion: other related issues

The concept of caring being someone’s identity came up in a few discussions – in some cases this was positive but in other cases negative. Some carers noted that they were viewed as having no status, or that people were not interested in their problems.

The concept of “loss” and missing out on opportunities that are available to other people was raised.

The relationship between caring and employment was discussed – sometimes being in work (out of the home) provided a break/protection from caring responsibilities, but the balancing work and care created significant challenges. In some cases, carers were unable to work due to their caregiving responsibilities. The financial impact of caring was also raised.

There were extensive discussions about the health and care systems and the support available to carers, with many expressing frustration and distress at the perceived deficiencies and lack of coordination of these systems.

Transitions into and out of caring were discussed, with some carers noting that becoming a carer was not a choice. There was a feeling that having a definite end to the caring role can be helpful. Participants also noted that caring responsibilities can change relatively rapidly, and the experience of caring may depend on where the cared for is living and the support available to them.

9 Appendix: Unpaid caring and EQ-5D

9.1 Background

In 2020, the National Institute for Health and Care Research Unit in Economic Methods of Evaluation of Health and Care Interventions (EEPRU) collected EQ-5D-3L and -5L response from a sample of 50,000 people in the UK. This dataset was used to estimate the relationship between EQ-5D-3L and 5L (Hernandez-Alava & Pudney, 2022). In addition to collecting demographic information, the survey also asked people if they provided care to anyone within or outside their household, and for how many hours per week. This presented an opportunity to analyse the relationship between different volumes of care and EQ-5D. When the UK value set becomes available for EQ-5D-5L (EuroQoL, 2023), I will analyse the effect of caring on 3L and 5L, to see how it differs. However, at the time of writing, the UK value set was not publicly available. I have therefore focussed on the effect of caring on EQ-5D-3L and explored the impact on 5L using the mapping to the 3L value set. I have also explored the responses to the different domains using 3L and 5L.

9.2 Methods

9.2.1 Data

The survey was conducted online using the OnePoll panel which was generally considered to be representative of the UK. Respondents were randomised to first complete either the 3L or the 5L, then the EQ-Visual Analogue Scale, answer to questions about age group, sex, family status, health, and caring, and finally to complete the 5L or 3L (whichever they had not answered first). A sample of 49,999 responses were included in the analysis. Full details can be found in the report (Hernandez Alava et al, 2020).

Respondents who answered that they provided care to someone living with or not living with them were asked how many hours of care they provided per week, with categories of 0-4, 5-9, 10-19, 20-34, 35-49, 50-99, 100+/continuous care and “don’t know”.

9.2.2 Analysis

For the EQ-5D-3L and 5L mapped to 3L index scores, I estimated the average treatment effect on the treated (ATT) where the “treatment” was providing unpaid care, at different weekly volumes. I considered as unpaid carers everyone who reported providing care either inside or outside the household, or both. I combined the two highest levels of care hours per week (50-99 and 100+) to increase the sample size. This is consistent with previous research which found the 20 or 50 hours/week thresholds to be important (Carers UK., 2023).

I considered potential confounders from the dataset: age group, sex, family status, education, and diagnosis of a health condition. I believed that these confounders could potentially affect both the likelihood of somebody providing unpaid care (at different volumes) and their EQ-5D. The Directed Acyclic Graph is shown in Figure 10.

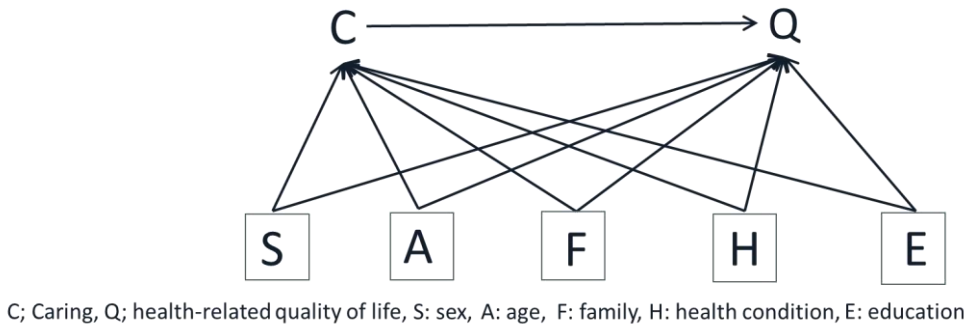


Figure 10: Directed Acyclic Graph

9.2.3 Inverse probability weighting (IPW)

I used a multinomial logit model to predict the probability of belonging to each class of weekly caregiving hours as a function of the confounders. The inverse of these probabilities is used to assign weights to the individuals in the sample. The goal is to reweight the observations so that the treatment and control groups become more comparable, giving lower weights to individuals who are over-represented in one group and under-represented in the other and vice versa.

For the purposes of modelling, I collapsed some of the categories for confounders. I collapsed the categories for family status to combine people who answered that they were either single, in a relationship, widowed or divorced to create a larger group who were not living with a partner. I combined people who answered that they were either married or cohabiting to create a group who were living with a partner. Within each of these groups I differentiated between people who did and did not have children. I felt that the distinction between people living with a partner and not living with a partner was important when considering how this might particularly affect the probability of caring for someone within the household (since people living alone cannot care for someone else within the household) and also caring outside the household. The presence of children may also influence whether someone is needed to provide care within the household, or whether they are able to provide care outside the household.

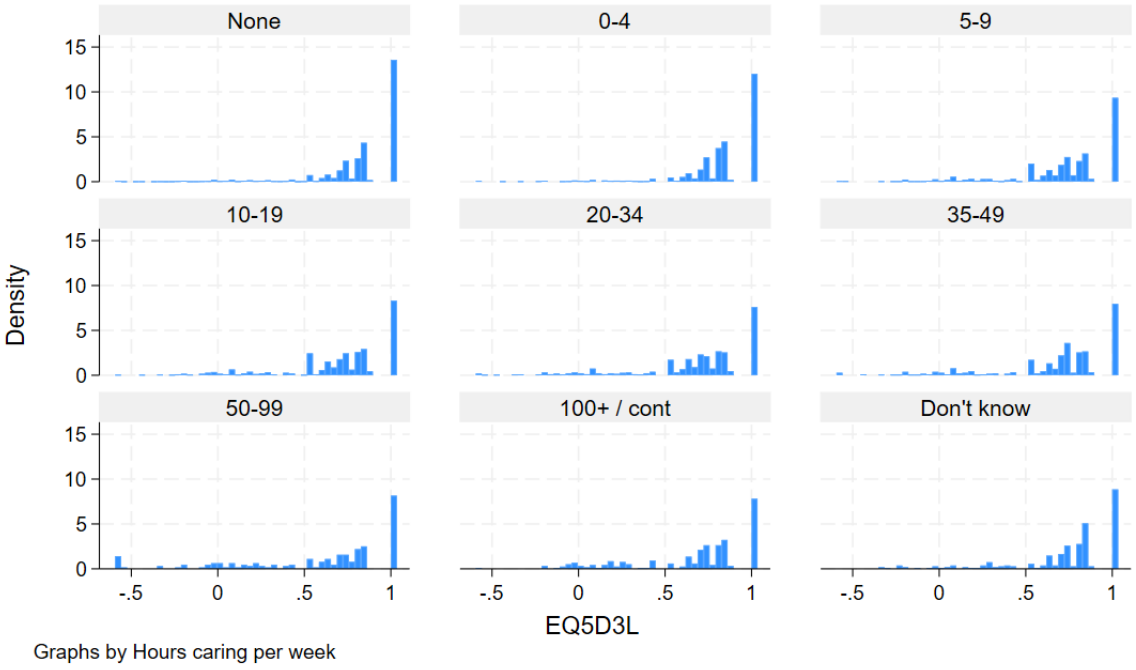
The education category “none of the above” had the fewest responses, so I combined this with “secondary education” to create a “secondary or below” category.

I included an age squared term to allow for a non-linear relationship with age.

9.2.4 Doubly robust methods

Regression adjustment (RA) is another technique to estimate the causal effect of a treatment while controlling for confounders. In RA, the outcome is regressed on the the confounders separately in the treatment and the control groups. Those regressions are then used to predict the counterfactuals for each individual in the sample, that is, the regression estimated in the treatment group is used to predict what the outcome would have been for those in the control group if they had been treated and vice versa. Thus, for each individual, we have both their observed outcome and their predicted counterfactual and their difference is an estimate of each individual's treatment effect. The causal effect of the treatment is simply the average across individuals of their estimated individual treatment effect. RA will not be appropriate where the regression model is not appropriate. The EQ-5D 3L index score (or 5L mapped to 3L) does not follow a typical distribution (such as normal), as illustrated in Figure 11, and so a linear regression model would not accurately predict the EQ-5D index score, and so the RA would not be appropriate. However, a probit model is appropriate for predicting the probability of full health or not.

When we combine regression adjustment with IPW, the approach is termed "doubly robust" since it requires either the IPW or the RA to be correctly specified, but not both. I used doubly robust methods for predicting the probability of reporting full health, but not for predicting EQ-5D index scores.



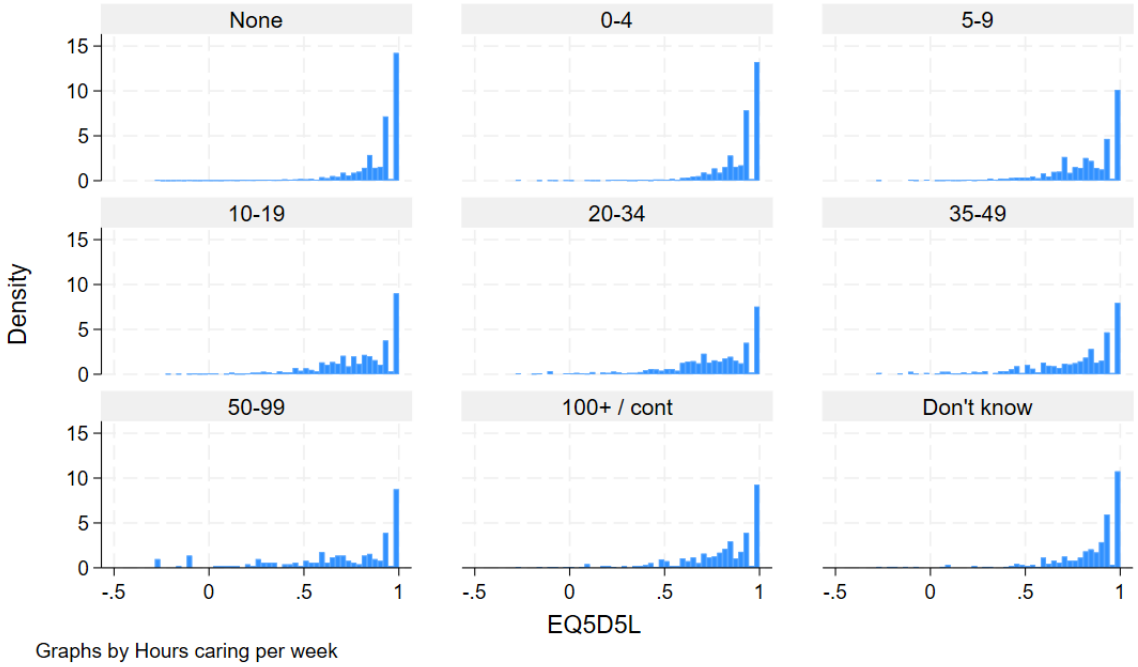


Figure 11: Distribution of EQ-5D-3L and 5L scores, by volume of care

I used an ordered probit model to estimate the effect of caring on responding at the different levels to each domain.

9.3 Results

Table 3 and Table 4 present baseline characteristics for the whole sample, and split for not-carers (75.4%), carers who care inside the home (7.2%), carers who care outside the home (10.3%), and carers who care both inside and outside the home (7.1%). Not-carers can be seen to have higher EQ-5D-3L and 5L scores than carers. Not-carers are more likely to be single without children and less likely to be married with children. The distribution of education and age across carer categories is less clear.

Table 3: Baseline characteristics (1)

	Carer status				
	Not carer	**Carer inside home**	**Carer outside home**	**Carer inside and outside home**	**Total**
N (%)	36,901 (75.4%)	3,536 (7.2%)	5,047 (10.3%)	3,451 (7.1%)	48,935 (100.0%)
Hours caring per week					
None	36,901 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	36,901 (86.2%)
0-4	0 (0.0%)	311 (18.4%)	1,175 (46.4%)	162 (9.6%)	1,648 (3.8%)
5-9	0 (0.0%)	252 (14.9%)	593 (23.4%)	348 (20.7%)	1,193 (2.8%)
10-19	0 (0.0%)	290 (17.1%)	327 (12.9%)	401 (23.8%)	1,018 (2.4%)
20-34	0 (0.0%)	230 (13.6%)	190 (7.5%)	304 (18.0%)	724 (1.7%)
35-49	0 (0.0%)	171 (10.1%)	132 (5.2%)	209 (12.4%)	512 (1.2%)
50-99	0 (0.0%)	66 (3.9%)	19 (0.8%)	96 (5.7%)	181 (0.4%)
100+ / cont	0 (0.0%)	189 (11.2%)	28 (1.1%)	121 (7.2%)	338 (0.8%)
Don't know	0 (0.0%)	185 (10.9%)	69 (2.7%)	44 (2.6%)	298 (0.7%)
Age group					
18-24	2,959 (8.0%)	319 (9.0%)	368 (7.3%)	332 (9.6%)	3,978 (8.1%)
25-34	6,457 (17.5%)	530 (15.0%)	716 (14.2%)	831 (24.1%)	8,534 (17.4%)
35-44	6,980 (18.9%)	632 (17.9%)	920 (18.2%)	993 (28.8%)	9,525 (19.5%)
45-54	6,064 (16.4%)	644 (18.2%)	1,144 (22.7%)	618 (17.9%)	8,470 (17.3%)
55-64	6,731 (18.2%)	644 (18.2%)	1,259 (24.9%)	417 (12.1%)	9,051 (18.5%)
65-74	6,104 (16.5%)	571 (16.1%)	568 (11.3%)	208 (6.0%)	7,451 (15.2%)
75+	1,606 (4.4%)	196 (5.5%)	72 (1.4%)	52 (1.5%)	1,926 (3.9%)
Sex					
Male	18,118 (49.1%)	1,720 (48.6%)	2,140 (42.4%)	1,842 (53.4%)	23,820 (48.7%)
Female	18,783 (50.9%)	1,816 (51.4%)	2,907 (57.6%)	1,609 (46.6%)	25,115 (51.3%)

Table 4: Baseline characteristics (2)

	Carer status				
	Not carer	**Carer inside home**	**Carer outside home**	**Carer inside and outside home**	**Total**
Family status					
Living alone + no children	10,850 (29.4%)	837 (23.7%)	1,264 (25.0%)	688 (19.9%)	13,639 (27.9%)
Living alone + children	5,713 (15.5%)	442 (12.5%)	833 (16.5%)	545 (15.8%)	7,533 (15.4%)
Living with partner + no children	5,468 (14.8%)	526 (14.9%)	675 (13.4%)	545 (15.8%)	7,214 (14.7%)
Living with partner + children	14,870 (40.3%)	1,731 (49.0%)	2,275 (45.1%)	1,673 (48.5%)	20,549 (42.0%)
Education level					
Secondary or below	8,357 (22.6%)	882 (24.9%)	941 (18.6%)	660 (19.1%)	10,840 (22.2%)
Post-Secondary	6,453 (17.5%)	651 (18.4%)	902 (17.9%)	535 (15.5%)	8,541 (17.5%)
Vocational	6,471 (17.5%)	722 (20.4%)	993 (19.7%)	575 (16.7%)	8,761 (17.9%)
Undergrad degree	9,844 (26.7%)	779 (22.0%)	1,342 (26.6%)	812 (23.5%)	12,777 (26.1%)
Postgrad degree	4,660 (12.6%)	391 (11.1%)	724 (14.3%)	662 (19.2%)	6,437 (13.2%)
Doctorate	1,116 (3.0%)	111 (3.1%)	145 (2.9%)	207 (6.0%)	1,579 (3.2%)
EQ5D3L	0.829 (0.239)	0.705 (0.297)	0.801 (0.234)	0.674 (0.333)	0.806 (0.256)
EQ5D5L	0.879 (0.178)	0.781 (0.227)	0.861 (0.169)	0.752 (0.252)	0.861 (0.191)

Table 3 and Table 4 demonstrate that there are some differences between the groups, but generally many similarities too, such that they do not look like completely different groups of individuals. IPW is therefore a good starting point to try and make the groups more comparable to estimate the causal effect, since there is already some overlap. The IPW model improved the level of balance for the included variables, with all weighted standardized differences close to zero and the variance ratios close to one (note that inference here is informal as there are no standard errors on these statistics(stata.com)).The overlap plot, shown in Figure 12 did not show much probability mass near 0 or 1 for any of

the care hour categories and the 8 classes have most of their respective masses in overlapping regions. Therefore, I considered that the overlap assumption is not violated(stata.com, 2025).

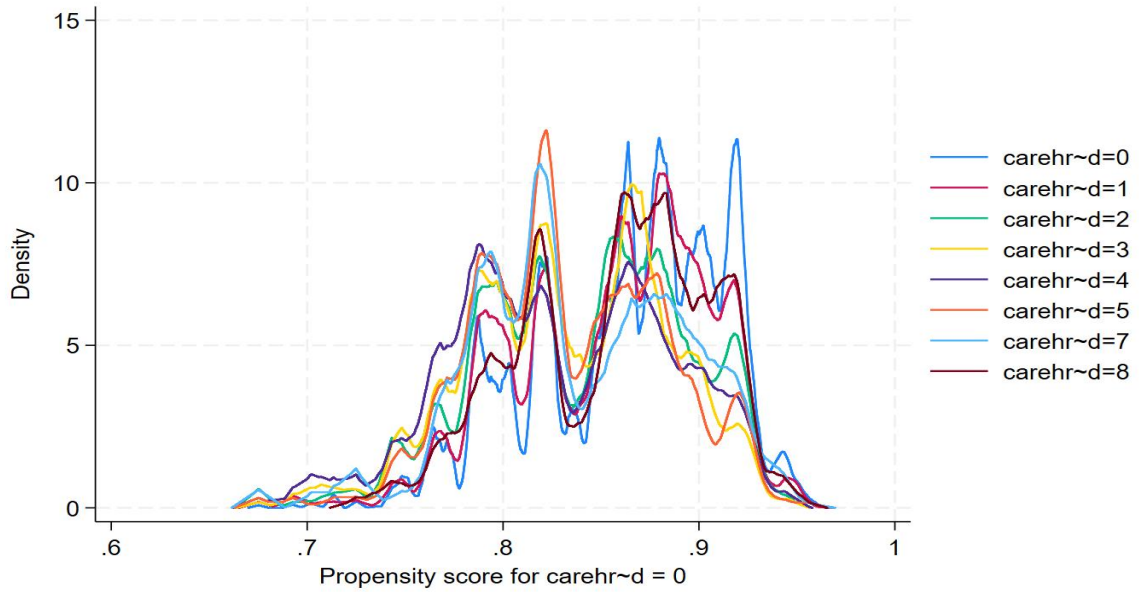


Figure 12: Overlap plot

The ATT relative to not caring, on the 3L index values are shown in Table 5, and 5L mapped to 3L in Table 6. The relative treatment effect was negative and statistically significant for all care categories with the exception of 0-4 hours per week for 3L or 5L and the “don’t know” category for 5. In general, higher hours of care were associated with larger HRQoL reductions, with the largest decrement in those providing 50+ hours/week and the smallest in the 5-9 hours group. The effect sizes were broadly similar for the 10-19, 20-34, and 35-49 hours/week categories, suggesting that caring does negatively affect HRQoL, particularly at very high volumes. The effect sizes were larger when using 3L than 5L.

Table 5: Average treatment effect on the treated: 3L index score

Care hours/week	**Coefficient**	**p-value**	**95% confidence interval lower bound**	**95% confidence interval upper bound**
Non-carer	0.826	<0.001	0.824	0.829
Average treatment effect on the treated relative to non-carer				
0-4	-0.001	0.920	-0.011	0.010
5-9	-0.072	<0.001	-0.087	-0.056
10-19	-0.097	<0.001	-0.118	-0.076
20-34	-0.099	<0.001	-0.123	-0.074
35-49	-0.090	<0.001	-0.119	-0.061
50+	-0.157	<0.001	-0.194	-0.121
Don't know	-0.049	<0.001	-0.076	-0.023

Table 6: Average treatment effect on the treated: 5L mapped to 3L index score

Care hours/week	**Coefficient**	**p-value**	**95% confidence interval lower bound**	**95% confidence interval upper bound**
Non-carer	0.877	<0.001	0.875	0.879
Average treatment effect on the treated relative to non-carer				
0-4	0.008	0.056	0.000	0.015
5-9	-0.044	<0.001	-0.055	-0.034
10-19	-0.071	<0.001	-0.086	-0.056
20-34	-0.082	<0.001	-0.101	-0.063
35-49	-0.076	<0.001	-0.097	-0.054
50+	-0.109	<0.001	-0.138	-0.081
Don't know	-0.022	0.044	-0.044	-0.001

The ATT on the probability of reporting full health is shown for 3L in Table 7 and 5L in Table 8. With exceptions of carers caring for 0-4 hours/week and those who didn't know how many hours they spent caring, carers were statistically significantly less likely to report full health than non-carers. The effect size was similar across hours of care but appeared largest for

people caring for 20-34 hours of care per week for both 3L and 5L. Non-carers were more likely to report full health with 3L than 5L, but effect sizes were larger for 3L than 5L, suggesting the reduction in the proportion of people reporting full health was smaller for 5L than 3L.

Table 7: Average treatment effect on the treated: 3L full health

Care hours/week	**Coefficient**	**p-value**	**95% confidence interval lower bound**	**95% confidence interval upper bound**
Non-carer	0.464	<0.001	0.459	0.469
Average treatment effect relative to non-carer				
0-4	-0.040	0.001	-0.064	-0.016
5-9	-0.115	<0.001	-0.143	-0.087
10-19	-0.137	<0.001	-0.169	-0.105
20-34	-0.165	<0.001	-0.204	-0.127
35-49	-0.104	<0.001	-0.150	-0.057
50+	-0.125	<0.001	-0.168	-0.081
Don't know	-0.102	0.001	-0.164	-0.039

Table 8: Average treatment effect on the treated: 5L full health

Care hours/week	**Coefficient**	**p-value**	**95% confidence interval lower bound**	**95% confidence interval upper bound**
Non-carer	0.392	<0.001	0.387	0.396
Average treatment effect relative to non-carer				
0-4	-0.012	0.313	-0.035	0.011
5-9	-0.091	<0.001	-0.118	-0.064
10-19	-0.098	<0.001	-0.129	-0.067
20-34	-0.133	<0.001	-0.170	-0.097
35-49	-0.082	<0.001	-0.127	-0.038
50+	-0.071	0.001	-0.113	-0.029
Don't know	-0.051	0.103	-0.112	0.010

Sankey charts demonstrate the change in responses to the anxiety/depression domain for the not-carers, and the carers 50+ hours/week group in Figure 13 and Figure 14. These illustrate how respondents' answers changed between 3L and 5L in different categories of care, rather than the relative change between groups. They are helpful in understanding how 3L and 5L responses differ, and so why the relative effect of caring may differ for 3L and 5L. They demonstrate that there was usually a proportion of people who reported "no problems" and "some problems" in 3L moving into "slight problems" with 5L, and generally a proportion of people who reported "extreme problems" with 3L moving into "severe problems" with 5L. Generally, non-carers were more likely to report "no problems" and carers more likely to report "extreme problems" across all domains. However, the Sankey diagrams do illustrate some inconsistencies in response across the 3L and 5L, for example for 50+hours carers answering about usual activities (Figure 15), there were people who reported "extreme problems" on 3L and "no problems" on 5L, and people who reported "no problems" on 3L and "severe problems" on 5L. This does call into question how valid the comparison of responses to 3L and 5L are, because it is unclear why people would give such dramatically different answers.

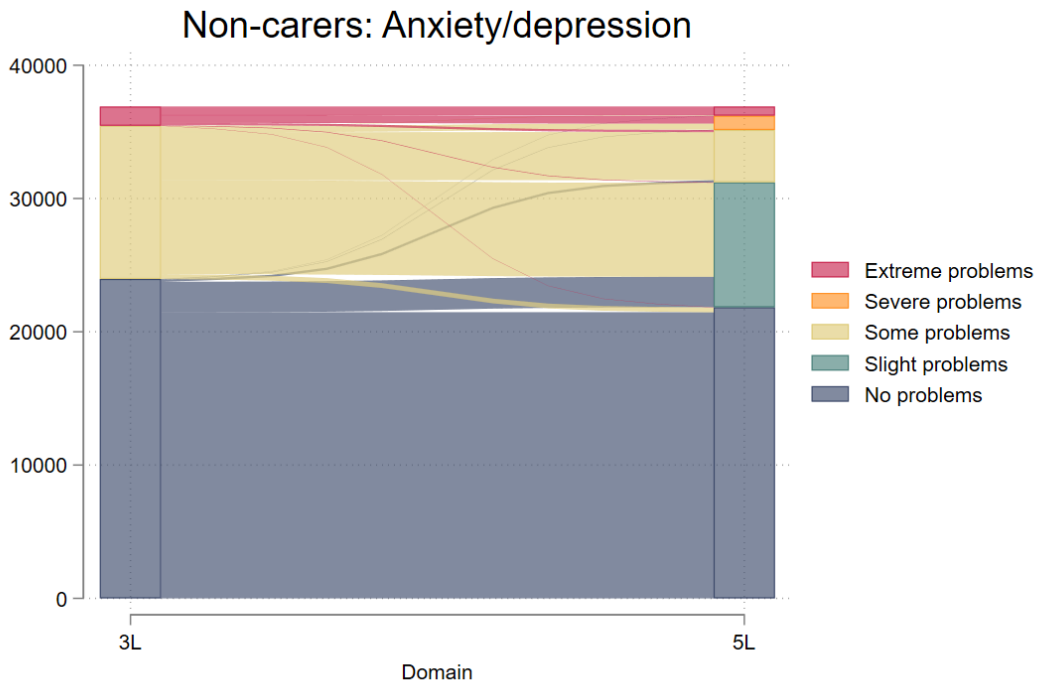


Figure 13: Sankey diagram: Anxiety/depression: Non-carers

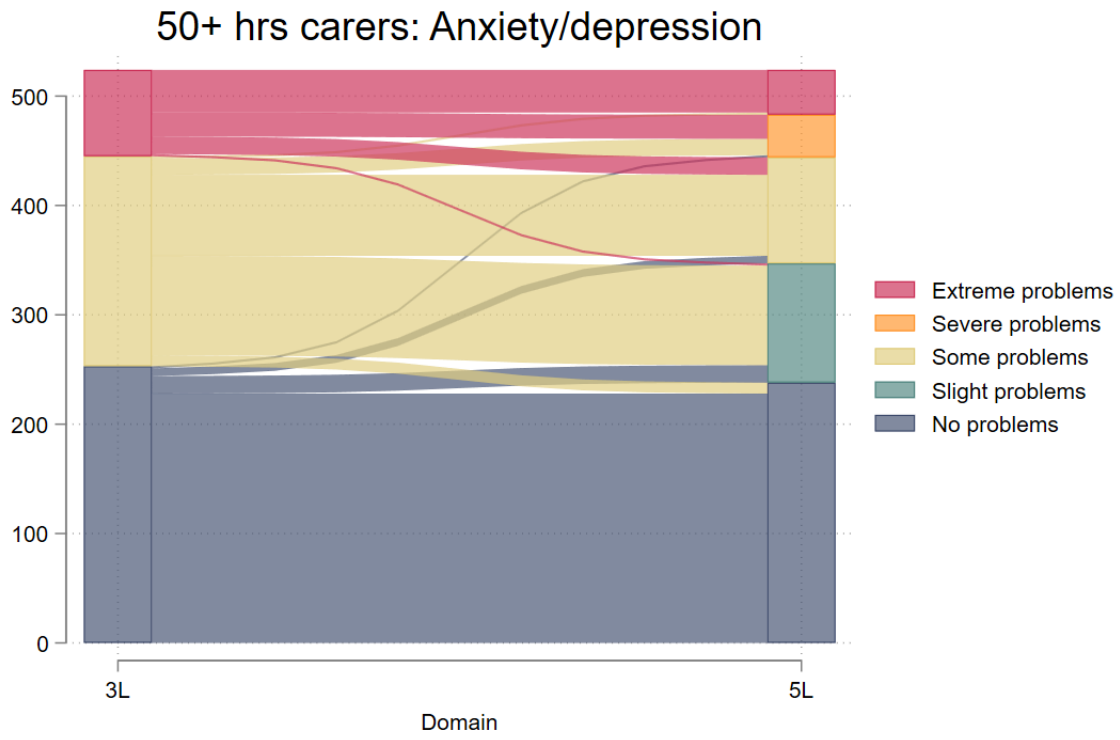


Figure 14: Sankey diagram: Anxiety/depression: 50+ hrs carers

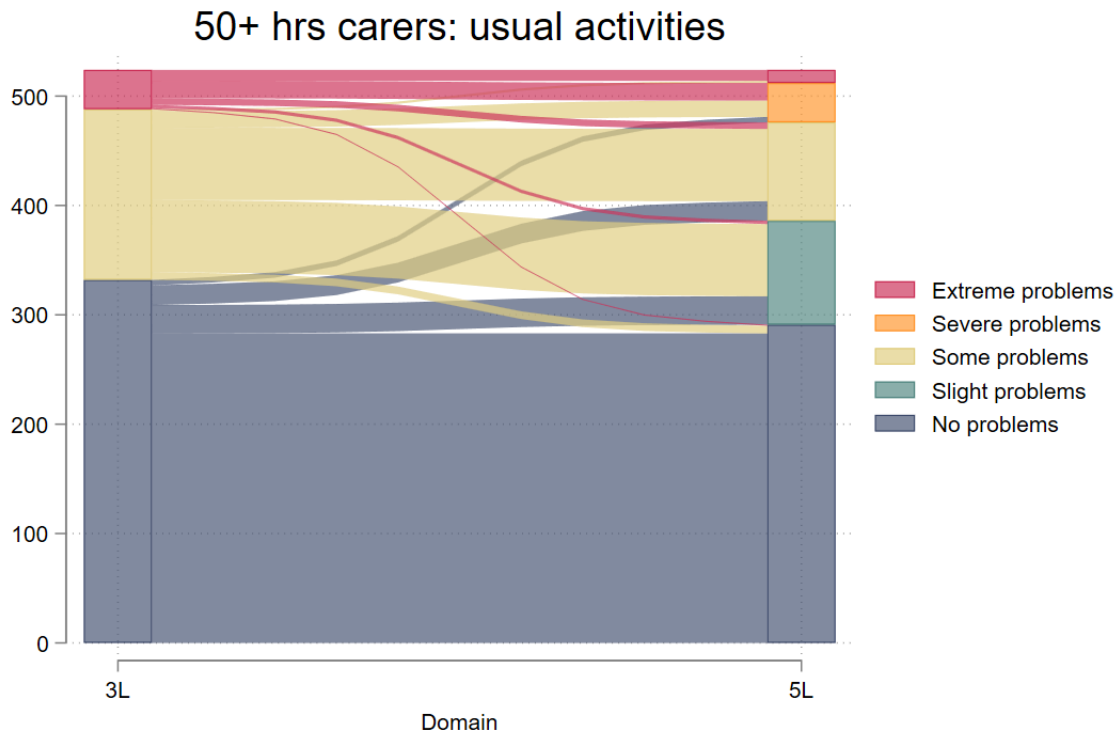


Figure 15: Sankey diagram: Usual activities: 50+ hours carers

Figure 16 presents the marginal effects for the carers 50+ hours/week group relative to not-carers, for the domains of 3L and 5L. Marginal effects quantify how a change in the independent variable (in this case, moving from 0 to 50+ hours/week), affect the predicted probability of the outcome (in this case, reporting each answer within a domain), holding all other variables constant. There were statistically significant effects across all domains for 3L and 5L, but effects for some specific individual levels were not statistically significant. Generally, carers were less likely to report “no problems” and more likely to report problems at some level. For mobility, self-care and usual activities, using 5L increased the reduction in respondents reporting no problems. On its own, we would expect that this would increase the average treatment effect for 5L compared to 3L. However, some of the carers who had reported the worst level on 3L (extreme problems) reported the second-worst level (severe problems) using 5L. This is particularly important because 3L has an “N3” term: a disutility which is applied to the index score if the respondent reports the worst level on any domain (it will be interesting to see if the 5L value set has a similar term). We tended to see this pattern across the domains for the different volumes of care although it was not always statistically significant. For pain/discomfort and anxiety/depression for the 50+ hours group, some of the non-carers who had reported “no problems” on 3L reported “slight problems” on 5L This decreased the difference between carers and non-carers and so further reduced the ATT for 5L. However, this finding is specific to these two domains for the 50+ hours group.

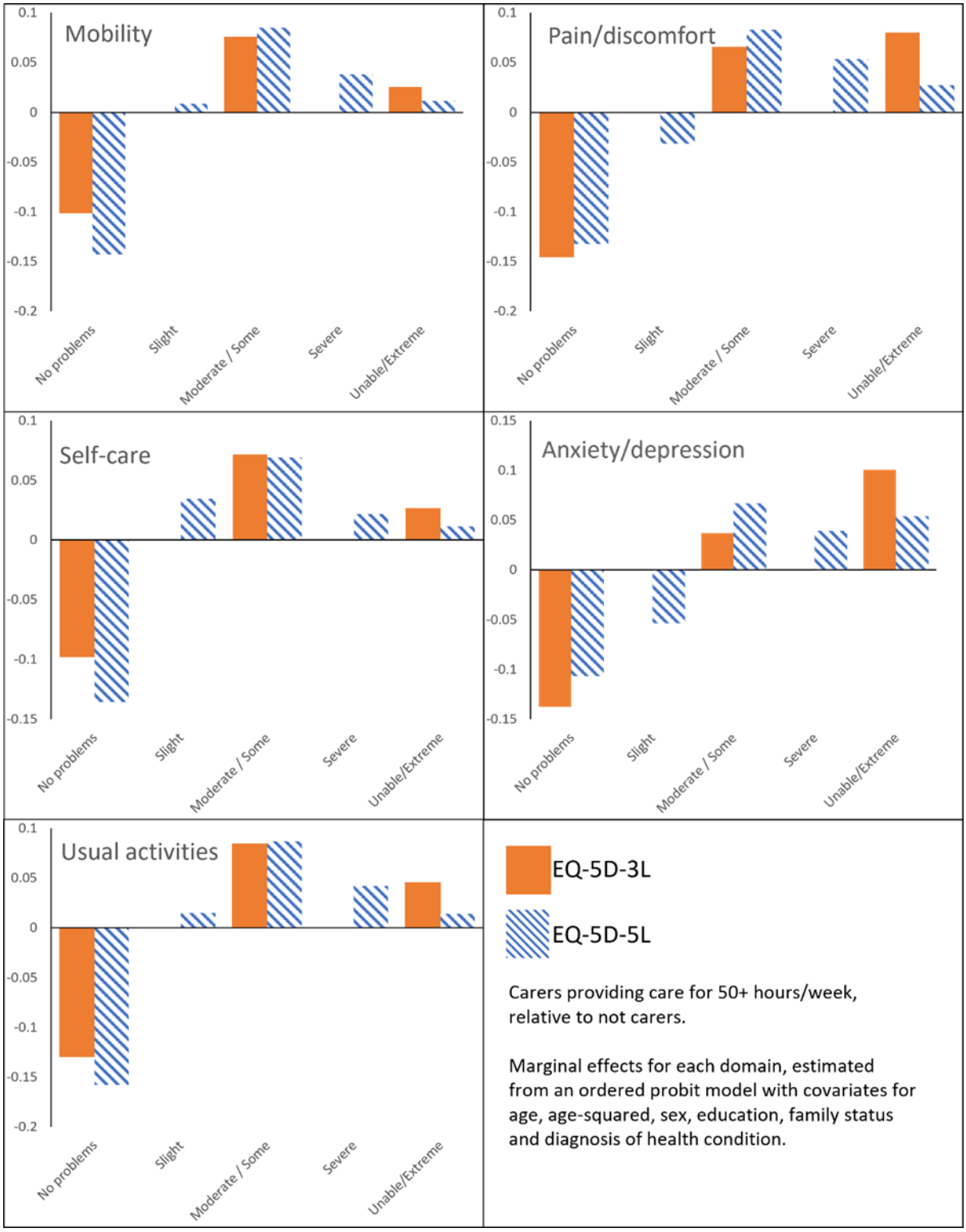


Figure 16: Marginal effects of domains, 50 hours/week+ group

Sensitivity analysis using the doubly robust estimand combining IPW and linear regression for 3L and 5L index scores found similar ATEs in Table 9 and Table 10. Whilst I do not

believe the linear regression is a good choice of model for the EQ5D distribution, this provided some reassurance that our estimates are robust.

Table 9: Average treatment effect on the treated: 3L index score, doubly robust

Care hours/week	**Coefficient**	**p-value**	**95% confidence interval lower bound**	**95% confidence interval upper bound**
Non-carer	0.826	<0.001	0.824	0.829
Average treatment effect on the treated relative to non-carer				
0-4	-0.001	0.920	-0.011	0.010
5-9	-0.072	<0.001	-0.087	-0.056
10-19	-0.097	<0.001	-0.118	-0.076
20-34	-0.099	<0.001	-0.123	-0.074
35-49	-0.090	<0.001	-0.119	-0.061
50+	-0.157	<0.001	-0.194	-0.121
Don't know	-0.049	<0.001	-0.076	-0.023

Table 10: Average treatment effect on the treated: 5L mapped to 3L index score, doubly robust

Care hours/week	**Coefficient**	**p-value**	**95% confidence interval lower bound**	**95% confidence interval upper bound**
Non-carer	0.877	<0.001	0.875	0.879
Average treatment effect on the treated relative to non-carer				
0-4	0.008	0.056	0.000	0.015
5-9	-0.044	<0.001	-0.055	-0.034
10-19	-0.071	<0.001	-0.086	-0.056
20-34	-0.082	<0.001	-0.101	-0.063
35-49	-0.076	<0.001	-0.097	-0.054
50+	-0.109	<0.001	-0.138	-0.081
Don't know	-0.022	0.044	-0.044	-0.001

In scenario analyses, I found the ATT was greater for both 3L and 5L when analysing the subpopulation of carers who provided care inside the home (excluding carers who only provide care outside the house, but including carers who provide care both inside and outside the home), but that the 3L average treatment effect was always bigger than for 5L (see Appendix Table 4). Effect sizes were statistically significant for 0-4 hours/week for 3L and 5L.

Table 11: Average treatment effect on the treated: 3L index score, exclude outside household carers

Care hours/week	**Coefficient**	**p-value**	**95% confidence interval lower bound**	**95% confidence interval upper bound**
Non-carer	0.829	0.001	0.827	0.831
Average treatment effect on the treated relative to non-carer				
0-4	-0.084	<0.001	-0.121	-0.047
5-9	-0.130	<0.001	-0.167	-0.093
10-19	-0.152	<0.001	-0.187	-0.116
20-34	-0.138	<0.001	-0.181	-0.094
35-49	-0.130	<0.001	-0.178	-0.082
50+	-0.192	<0.001	-0.251	-0.132
Don't know	-0.079	<0.001	-0.121	-0.036

Table 12: Average treatment effect on the treated: 5L mapped to 3L index score, exclude outside household carers

Care hours/week	**Coefficient**	**p-value**	**95% confidence interval lower bound**	**95% confidence interval upper bound**
Non-carer	0.879	<0.001	0.877	0.881
Average treatment effect on the treated relative to non-carer				
0-4	-0.051	0.056	-0.078	-0.024
5-9	-0.095	<0.001	-0.119	-0.070
10-19	-0.117	<0.001	-0.144	-0.090
20-34	-0.087	<0.001	-0.118	-0.056
35-49	-0.103	<0.001	-0.139	-0.067
50+	-0.134	<0.001	-0.183	-0.085
Don't know	-0.031	0.044	-0.057	-0.006

When I excluded inside-household carers, there were very few people in older age groups who provided care outside the house, so I collapsed the categories such that the greatest volume of care was 35+hours/week, and in this case the effect sizes were smaller than for the combined carer population of the inside-household carers.

Table 13: Average treatment effect on the treated: 3L index score, exclude inside household carers

Care hours/week	Coefficient	p-value	95% confidence interval lower bound	95% confidence interval upper bound
Non-carer	0.829	<0.001	0.827	0.831
Average treatment effect on the treated relative to non-carer				
0-4	0.019	0.001	0.007	0.030
5-9	-0.046	<0.001	-0.068	-0.025
10-19	-0.071	<0.001	-0.103	-0.038
20-34	-0.125	<0.001	-0.174	-0.076
35+	-0.032	0.067	-0.067	0.002
Don't know	-0.028	0.376	-0.091	0.035

Table 14: Average treatment effect on the treated: 5L mapped to 3L index score, exclude inside household carers

Care hours/week	Coefficient	p-value	95% confidence interval lower bound	95% confidence interval upper bound
Non-carer	0.879	<0.001	0.877	0.881
Average treatment effect on the treated relative to non-carer				
0-4	0.022	<0.001	0.014	0.031
5-9	-0.022	0.003	-0.036	-0.008
10-19	-0.048	<0.001	-0.070	-0.025
20-34	-0.091	<0.001	-0.128	-0.054
35+	-0.045	0.008	-0.079	-0.012
Don't know	-0.005	0.809	-0.045	0.035

9.4 Discussion

My analysis demonstrated that people who provide unpaid care for 5 or more hours per week have statistically significantly worse HRQoL, as measured by EQ-5D 3L, than people who do not provide care. This was true even after adjusting for confounding factors,

suggesting there may be a causal link between caring and HRQoL. Providing higher volumes of care generally had a bigger detrimental effect on HRQoL, with a maximum ATT of -0.157 for people providing 50+ hours of care per week. Within the same weekly hours of care, the ATT was larger for people providing care within the household than those only providing care outside the household. This suggests that co-resident carers experience a particularly high HRQoL detriment.

My analysis indicates that an increased volume of caregiving affects all domains of EQ-5D. It is unsurprising that caregiving affects anxiety/depression, and feasible that caregiving impedes usual activities due to time restrictions. It is less clear why caregiving would also impact pain, mobility and self-care, unless it has directly caused a physical injury (for example, back pain as a result of lifting the care recipient). It is therefore possible that there are differences between carers and non-carers (and carers at different volumes) in my sample that are not entirely controlled for. Therefore, while it is true that carers have worse HRQoL than non-carers, the differences in HRQoL may not be entirely due to caregiving.

The ATT was smaller for 5L mapped to 3L index score than the 3L index score. This is consistent with evidence for EQ-5D in patients, where 5L shifts all scores upwards and into a smaller range. It remains to be seen whether this holds true with the new UK 5L value set, but if so, it would mean that interventions targeted at carers are less cost-effective using 5L than 3L. It would also mean that the carer spillover impacts from patient interventions are smaller using 5L than 3L, but the impact on cost-effectiveness will depend on whether the intervention affects mortality and the method used to model carer HRQoL (since the carer HRQoL gain due to improved patient HRQoL would be smaller, but any carer QALY loss due to the “carer QALY trap” would also be smaller).

The size of the ATT for caring is fairly consistent with the range reported in the literature for EQ-5D: a study of 910 adult/elderly family carers in Sweden found a disutility of 0.015 (Davidson et al., 2008) and a study of 306 mothers of children with congenital abnormalities found a disutility of 0.10(Poley et al., 2012). An analysis of the Irish EQ-5D-5L survey found a total disutility from people caring for others with a serious illness was 0.014, with a statistically significant effect only on the anxiety/depression domain.

This was a large study with over 17,000 carers and nearly 37,000 non-carers for comparison. The sample size and variables included meant that I was able to exclude confounder bias from important confounders using causal inference methodology. There was good overlap between groups and the results appear robust to sensitivity analysis. This analysis provides data on the effect of caring on EQ-5D by volume of care in the UK. It is therefore a useful source for informing economic evaluation in the UK, since NICE states a

preference for the EQ-5D (National Institute for Health and Care Excellence, 2022) and there is a gradient of disutility by care volume which could be generalised to patient severity states in economic models.

There are a few limitations that stem from the study design. Firstly, the data about the carer is fairly limited: while I controlled for all of the confounders that we have information on, I did not have details on the caregivers' health conditions, or information such as duration of caregiving which previous research has indicated may be important(Pennington et al., 2025b). Secondly, the data is cross-sectional in nature so I could not analyse the change in HRQoL due to caring. I aimed to mitigate this through the doubly robust approach, but longitudinal data would be preferable to identify a causal effect. Lastly, I did not have any information about the care-recipient, such as their age and relationship with the carer, or their health condition/HRQoL. This is also likely to influence the carers' HRQoL(Bobinac et al., 2010, Bobinac et al., 2011), whereas this analysis assumes that the volume of caregiving fully captures all effects.

10 Appendix: Public and patient involvement and engagement: a new method for modelling carer's HRQoL

10.1 Introduction

10.1.1 Set-up

In April 2025 I organised a workshop to speak to people with experience of unpaid/informal caring to discuss how carers' HRQoL has been included in NICE appraisals so far, and to get feedback on my proposed method. The workshop was conducted on Google Meet and lasted 2.5 hours.

10.1.2 Participants

The workshop was attended by nine people who were either carers themselves or represented carers organisations (or both) and had experience of NICE appraisals. There were two members of NICE's People and Communities team, acting independently, whose role was to facilitate small group discussions and take notes.

10.2 Discussion: NICE decision-making and how carers are included

Generally, people felt that it was unclear to what extent NICE considered the impact of conditions on carers in their decision-making because the information discussed in public is often complex, and because decision-making is conducted in private. There were concerns that patient evidence, and particularly qualitative evidence, is not given the same importance as other, more quantitative data. Other people felt that qualitative evidence had been persuasive in some NICE appraisals. Attendees felt that qualitative evidence is important for capturing carers' experiences, and that it can feel like somebody else is scoring carer's HRQoL. There were some suggestions that carers and patients should have more involvement in NICE decisions, but concerns were raised about the resources this would require for patient and carer organisations.

10.3 Discussion: Utility and disutility approach (current methods)

Attendees were generally familiar with the disutility and utility approaches used in NICE appraisals, but some noted the disutility approach was more familiar. People felt that each approach had strengths and limitations, making it difficult to decide if one was better than the other, and that that neither approach fully captured the impact of caring or really considered the relationship between patients and carers. There was a sentiment that the evidence assessment group (EAG) always counter the company's approach, but that neither approach

is sufficiently challenged. There was some reluctance to use a “one size fits all” approach when conditions differ.

10.4 Discussion: Identifying the impact of caring on HRQoL

I introduced the idea of measuring the impact of caring on HRQoL, and how this would differ from either measuring carer utilities or carer disutilities.

Attendees discussed that it can be challenging to identify who carers are, and that consideration should be given to the wider family (for example siblings) and not just the primary carer. The challenges of quantifying the impact of caring were discussed, with people noting that for some carers this is 24 hours a day, and that it can be hard to separate caring responsibilities from responsibilities associated with the relationship of the carer (for example as a parent). There was a feeling that while it is encouraging that researchers are trying to address this, even improved methods and measures would not capture all impacts of caring.

There was some recognition that NICE needs a consistent approach, but the importance of flexibility and the variability between conditions and populations was highlighted, with a suggestion that discussion at scoping stage could allow for a bespoke approach.

10.5 Discussion: Using the family effect and caregiving effect in NICE appraisals

I introduced the concepts of the family effect and caregiving effect, gave an example of an analysis where these were calculated, and explained how the two effects could be combined in modelling carers’ HRQoL.

There was some support for an approach that bridged the gap between the utility and disutility approaches, but concerns about the practicalities of separating the two effects for each appraisal. Attendees understood the two concepts but felt that the terminology was unclear, with the “family effect” expected to capture the whole family. There was support for recognising the family effect, and a feeling that should be extended to the wider family and not just the primary carer. There were some concerns about how this approach would capture the effects of bereavement and whether it captured the positive aspects of caring.

There were some concerns about the implications of the approach, and in particular whether including only the effect of caring would be less favourable than including utilities (the utility approach). People noted the need to be clear about what is being captured and measured so that it does not unfairly penalise carers.

The variability between conditions was discussed, with attendees noting that the impact of caring will differ between conditions and that length of life may be particularly important in some cases. It was suggested that the approach might work well for rare diseases where there is a lot of uncertainty.

11 References

Al-Janabi, H., Manca, A. & Coast, J. (2017) Predicting carer health effects for use in economic evaluation. *PLoS One*, 12(9), e0184886.

Al-Janabi, H., McLoughlin, C., Oyebode, J., Efstathiou, N. & Calvert, M. (2019) Six mechanisms behind carer wellbeing effects: A qualitative study of healthcare delivery. *Soc Sci Med*, 235, 112382.

Al-Janabi, H., Nicholls, J. & Oyebode, J. (2016a) The need to "carer proof" healthcare decisions. *BMJ*, 352, i1651.

Al-Janabi, H., van Exel, J., Brouwer, W. & Coast, J. (2016b) A Framework for Including Family Health Spillovers in Economic Evaluation. *Med Decis Making*, 36(2), 176-86.

Al-Janabi, H., Van Exel, J., Brouwer, W., Trotter, C., Glennie, L., Hannigan, L. & Coast, J. (2016c) Measuring Health Spillovers for Economic Evaluation: A Case Study in Meningitis. *Health Econ*, 25(12), 1529-1544.

Al-Janabi, H., Wittenberg, E., Donaldson, C. & Brouwer, W. (2022) The relative value of carer and patient quality of life: A person trade-off (PTO) study. *Soc Sci Med*, 292, 114556.

Al-Rabayah, A. A., Al Froukh, R. F., Al Najjar, B., Rayyan, M., Salmany, S., Iweir, S. & Nazer, L. (2022) Quality of Life of Family Caregivers of Critically Ill Patients With Cancer Before and After Intensive Care Unit Admission Measured by EQ-5D 3-Level: A Longitudinal Prospective Cohort Study. *Value Health Reg Issues*, 30, 39-47.

Angelini, V. & Costa-Font, J. (2023) Health and wellbeing spillovers of a partner's cancer diagnosis. *Journal of Economic Behavior & Organization*, 212, 422-437.

Basu, A. & Meltzer, D. (2005) Implications of spillover effects within the family for medical cost-effectiveness analysis. *Journal of Health Economics*, 24(4), 751-773.

Bates, M. (1989) The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5), 407-424.

Bertogg, A. P., P.; Raiber, K.; (2025) *Dynamics of Later-Life Caregiving and Health. Insights From Biomarker Data and Cognitive Tests.*, 2025. Available online: https://osf.io/preprints/socarxiv/e456k_v2 [Accessed.

Bhadhuri, A., Jowett, S., Jolly, K. & Al-Janabi, H. (2017) A Comparison of the Validity and Responsiveness of the EQ-5D-5L and SF-6D for Measuring Health Spillovers: A Study of the Family Impact of Meningitis. *Med Decis Making*, 37(8), 882-893.

Bobinac, A., van Exel, N. J., Rutten, F. F. & Brouwer, W. B. (2010) Caring for and caring about: disentangling the caregiver effect and the family effect. *J Health Econ*, 29(4), 549-56.

Bobinac, A., van Exel, N. J., Rutten, F. F. & Brouwer, W. B. (2011) Health effects in significant others: separating family and care-giving effects. *Med Decis Making*, 31(2), 292-8.

Bom, J., Bakx, P., Schut, F. & van Doorslaer, E. (2019) Health effects of caring for and about parents and spouses. *The Journal of the Economics of Ageing*, 14.

Bom, J. & Stockel, J. (2021) Is the grass greener on the other side? The health impact of providing informal care in the UK and the Netherlands. *Soc Sci Med*, 269, 113562.

Booth, A. S., A.; Papaioannou, D. (2016) *Systematic Approaches to a Succesful Literature Review*. London: SAFE.

Brouwer, W. B. F. (2019) The Inclusion of Spillover Effects in Economic Evaluations: Not an Optional Extra. *Pharmacoeconomics*, 37(4), 451-456.

Brown, R. M. & Brown, S. L. (2014) Informal Caregiving: A Reappraisal of Effects on Caregivers. *Social Issues and Policy Review*, 8(1), 74-102.

Callaway, B. & Sant'Anna, P. H. C. (2021) Difference-in-Differences with multiple time periods. *Journal of Econometrics*, 225(2), 200-230.

Carers UK. (2023) *State of Caring 2023. The impact of caring on: health.* , 2023. Available online: https://www.carersuk.org/reports/state-of-caring-survey-2023-the-impact-of-caring-on-health/ [Accessed.

Carers UK. (2024) *State of Caring.*, 2024. Available online: https://www.carersuk.org/policy-and-research/state-of-caring-survey/ [Accessed.

Charlton, V. (2022) Justice, Transparency and the Guiding Principles of the UK's National Institute for Health and Care Excellence. *Health Care Analysis*, 30(2), 115-145.

Cheng, M., Yang, H. & Yu, Q. (2023) Impact of informal caregiving on caregivers' subjective well-being in China: a longitudinal study. *Arch Public Health*, 81(1), 209.

Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., Devlin, N., Smith, P. C. & Sculpher, M. (2015) Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol Assess*, 19(14).

Cookson, R. (2015) Justice and the NICE approach. *Journal of Medical Ethics*, 41(1), 99.

Costa-Font, J. & Vilaplana-Prieto, C. (2022) Mental health effects of caregivers respite: Subsidies or Supports? *The Journal of the Economics of Ageing*, 23.

Dawoud, D., Lamb, A., Moore, A., Bregman, C., Rupniewska, E., Paling, T., Wolfram, V., Lovett, R. E. S. & Dent, R. (2022) Capturing what matters: updating NICE methods guidance on measuring and valuing health. *Quality of Life Research*, 31(7), 2167-2173.

Eibich, P. (2023) Instrumental variable estimates of the burden of parental caregiving. *The Journal of the Economics of Ageing*, 26, 100467.

Ervin, J., Taouk, Y., Fleitas Alfonzo, L., Peasgood, T. & King, T. (2022) Longitudinal association between informal unpaid caregiving and mental health amongst working age adults in high-income OECD countries: A systematic review. *eClinicalMedicine*, 53.

EuroQoL (2023) *New UK EQ-5D-5L valuation study blog*, 2023. Available online: https://euroqol.org/about-us/blogs/ [Accessed.

EuroQol (2025) *EQ-HWB.*, 2025. Available online: https://euroqol.org/research-at-euroqol/eq-hwb/ [Accessed.

Fisher, B. T., J. (1990) Towards a Feminist Theory of Care, in Abel, E. N., M. (ed), *Circles of Care.* Albany, NY.: SUNY Press., 36-54.

Folbre, N. (2001) *The invisible heart.* New York, Unitet States.: The New Press.

Froelich, L., Llado, A., Khandker, R. K., Pedros, M., Black, C. M., Sanchez Diaz, E. J., Chekani, F. & Ambegaonkar, B. (2021) Quality of Life and Caregiver Burden of Alzheimer's Disease Among Community Dwelling Patients in Europe: Variation by Disease Severity and Progression. *J Alzheimers Dis Rep*, 5(1), 791-804.

Frye, M. (1983) *The politics of reality: essays in feminist theory.* Berkeley.: Crossing Press.

Gilligan, C. (1993) *In a different voice.* Cambridge, Massachusetts. London, England.: Harvard University Press.

Grimm, S. E., Witlox, W., Wolff, R., Chalker, A., Hiligsmann, M., Wijnen, B., Ahmadu, C., Ryder, S., Armstrong, N., Duffy, S., Syndikus, I., Kleijnen, J. & Joore, M. A. (2022) Mogamulizumab for Previously Treated Mycosis Fungoides and Sezary Syndrome: An Evidence Review Group Perspective of a NICE Single Technology Appraisal. *Pharmacoeconomics*, 40(5), 509-518.

Han, S. H. (2023) Revisiting the caregiver stress process: Does family caregiving really lead to worse mental health outcomes? *Adv Life Course Res*, 58, 100579.

Han, S. H., Kim, K. & Burr, J. A. (2021) Take a sad song and make it better: Spousal activity limitations, caregiving, and depressive symptoms among couples. *Soc Sci Med*, 281, 114081.

Handels, R., Herring, W. L., Grimm, S., Skoldunger, A., Winblad, B., Wimo, A. & Jonsson, L. (2024) New International Pharmaco-Economic Collaboration on Alzheimer's Disease (IPECAD) Open-Source Model Framework for the Health Technology Assessment of Early Alzheimer's Disease Treatment: Development and Use Cases. *Value Health*.

Henry, E., Al-Janabi, H., Brouwer, W., Cullinan, J., Engel, L., Griffin, S., Hulme, C., Kingkaew, P., Lloyd, A., Payakachat, N., Pennington, B., Pena-Longobardo, L. M., Prosser, L. A., Shah, K., Ungar, W. J., Wilkinson, T. & Wittenberg, E. (2024) Recommendations for Emerging Good Practice and Future Research in Relation to Family and Caregiver Health Spillovers in Health Economic Evaluations: A Report of the SHEER Task Force. *Pharmacoeconomics*, 42(3), 343-362.

Henry, E. & Cullinan, J. (2024) Maternal Mental Health Spillovers From Child Illness and Disability: A Dynamic Panel Analysis. *Value Health*.

Hernandez-Alava, M. & Pudney, S. (2022) Mapping between EQ-5D-3L and EQ-5D-5L: A survey experiment on the validity of multi-instrument data. *Health Econ*, 31(6), 923-939.

Hernandez Alava, M., Pudney, S. & Wailoo, A. (2020) *Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from an English population study.*, 2020. Available online: https://doi.org/10.15131/shef.data.25219157.v1 [Accessed.

Hernández Alava, M., Pudney, S. & Wailoo, A. (2022) *Estimating EQ-5D by Age and Sex for the UK. NICE DSU Report.* , 2022. Available online: https://www.sheffield.ac.uk/nice-dsu/methods-development/estimating-eq-5d [Accessed.

Hughes, D. A., Bourke, S., Jones, A., Bhatt, R., Huda, S., Mutch, K. & Jacob, A. (2022) Health utilities and costs for neuromyelitis optica spectrum disorder. *Orphanet J Rare Dis*, 17(1), 159.

Jain, U. & Ma, M. (2024) Together in sickness and in health: Spillover of physical, mental, and cognitive health among older English couples. *Health Econ*, 33(9), 1989-2012.

Kanters, T. A., van Hezik-Wester, V., Boateng, A., Cranmer, H., Kvamme, I., Santi, I., Al-Janabi, H. & van Exel, J. (2024) Including carer health-related quality of life in

NICE health technology assessments in the United Kingdom. *Health Econ Policy Law*, 1-13.

Kirvalidze, M., Beridze, G., Wimo, A., Morin, L. & Calderon-Larranaga, A. (2023) Variability in perceived burden and health trajectories among older caregivers: a population-based study in Sweden. *J Epidemiol Community Health*, 77(2), 125-132.

Kittay, E. F. (2001) When Caring Is Just and Justice is Caring: Justice and Mental Retardation. *Public Culture*, 13(3), 557-580.

Kramer, M. D. & Bleidorn, W. (2024) The Well-Being Costs of Informal Caregiving. *Psychol Sci*, 35(12), 1382-1394.

Kröger, T. (2009) Care research and disability studies: Nothing in common? *Critical Social Policy*, 29(3), 398-420.

Labbas, E. & Stanfors, M. (2023) Does Caring for Parents Take Its Toll? Gender Differences in Caregiving Intensity, Coresidence, and Psychological Well-Being Across Europe. *Eur J Popul*, 39(1), 18.

Lamsal, R., Yeh, E. A., Pullenayegum, E. & Ungar, W. J. (2024) A Systematic Review of Methods and Practice for Integrating Maternal, Fetal, and Child Health Outcomes, and Family Spillover Effects into Cost-Utility Analyses. *Pharmacoeconomics*, 42(8), 843-863.

Landfeldt, E. & Sandhu, H. (2024) Economic Evaluations of Treatments for Duchenne Muscular Dystrophy: The Caregiver QALY Trap. *Pharmacoeconomics*, 42(5), 475-478.

Lavelle, T. A., Riley, A., Kowal, S., Shapouri, S., Fox, D., Zhang, H., Vesel, T., Belter, L., McCarthy O'Toole, C., Mulé, C. M. & Lerner, D. (2025) Mental Health, Quality of Life, and Work Functioning of Parents of Children Who Have Died of Spinal Muscular Atrophy. *Value Health*.

Le, D. D. & Ibuka, Y. (2023) Understanding the effects of informal caregiving on health and well-being: Heterogeneity and mechanisms. *Soc Sci Med*, 317, 115630.

Lee, D., Kim, Y. & Devine, B. (2022) Spillover Effects of Mental Health Disorders on Family Members' Health-Related Quality of Life: Evidence from a US Sample. *Med Decis Making*, 42(1), 80-93.

Lomas, J., Martin, S. & Claxton, K. (2019) Estimating the Marginal Productivity of the English National Health Service From 2003 to 2012. *Value Health*, 22(9), 995-1002.

Loveman, E. G., C. Kirby, J. Takeda, A. Picot, J. Payne, E. Clegg, A. (2006) The clinical and cost-effectiveness of donepezil, rivastigmine, galantamine and memantine for Alzheimer's disease. *Health Technology Assessment*, 10(1).

Majmudar, I. K., Mihalopoulos, C., Abimanyi-Ochom, J., Mohebbi, M. & Engel, L. (2024) The association between loneliness with health service use and quality of life among informal carers in Australia. *Soc Sci Med*, 348, 116821.

McCabe, C. (2019) Expanding the Scope of Costs and Benefits for Economic Evaluations in Health: Some Words of Caution. *Pharmacoeconomics*, 37(4), 457-460.

McEwen, B. S. & Stellar, E. (1993) Stress and the Individual: Mechanisms Leading to Disease. *Archives of Internal Medicine*, 153(18), 2093-2101.

McLoughlin, C., Goranitis, I. & Al-Janabi, H. (2020) Validity and Responsiveness of Preference-Based Quality-of-Life Measures in Informal Carers: A Comparison of 5 Measures Across 4 Conditions. *Value Health*, 23(6), 782-790.

Methi, F., Nes, R. B., Skirbekk, V. & Hansen, T. (2024) The double-edged sword of becoming a caregiver: dynamic impact on four dimensions of well-being in Norway. *BMC Psychol*, 12(1), 120.

Mott, D. J., Schirrmacher, H., Al-Janabi, H., Guest, S., Pennington, B., Scheuer, N., Shah, K. K. & Skedgel, C. (2023) Modelling Spillover Effects on Informal Carers: The Carer QALY Trap. *Pharmacoeconomics*, 41(12), 1557-1561.

Mulders-Manders, C. M., Kanters, T. A., van Daele, P. L. A., Hoppenreijs, E., Legger, G. E., van Laar, J. A. M., Simon, A. & Hakkaart-van Roijen, L. (2018) Decreased quality of life and societal impact of cryopyrin-associated periodic syndrome treated with canakinumab: a questionnaire based cohort study. *Orphanet J Rare Dis*, 13(1), 59.

Musters, S. C. W., Kreca, S. M., van Dieren, S., van der Wal-Huisman, H., Romijn, J. A., Chaboyer, W., Nieveen van Dijkum, E. J. M., Eskes, A. M. & Consortium, A. (2024) Family caregiver outcomes after participating in a hospital-based family involvement program after major gastrointestinal surgery: a subgroup analysis of a patient preferred cohort study. *Int J Surg*, 110(8), 4746-4753.

Naci, H., Murphy, P., Woods, B., Lomas, J., Wei, J. & Papanicolas, I. (2025) Population-health impact of new drugs recommended by the National Institute for Health and Care Excellence in England during 2000-20: a retrospective analysis. *Lancet*, 405(10472), 50-60.

National Institute for Health and Care Excellence (2018a) *Ocrelizumab for treating relapsing–remitting multiple sclerosis*, 2018a. Available online: https://www.nice.org.uk/guidance/ta533 [Accessed.

National Institute for Health and Care Excellence (2018b) *Strimvelis for treating adenosine deaminase deficiency–severe combined immunodeficiency.*, 2018b. Available online: https://www.nice.org.uk/guidance/hst7 [Accessed.

National Institute for Health and Care Excellence (2019a) *Patisiran for treating hereditary transthyretin amyloidosis*, 2019a. Available online: https://www.nice.org.uk/guidance/hst10/chapter/4-Consideration-of-the-evidence [Accessed.

National Institute for Health and Care Excellence (2019b) *Position statement on use of the EQ-5D-5L value set for England.*, 2019b. Available online: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l [Accessed.

National Institute for Health and Care Excellence (2021) *Risdiplam for treating spinal muscular atrophy*, 2021. Available online: https://www.nice.org.uk/guidance/ta755 [Accessed.

National Institute for Health and Care Excellence (2022) *NICE health technology evaluations: the manual.*, 2022. Available online: https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741 [Accessed.

National Institute for Health and Care Excellence. (2023a) *Ataluren for treating Duchenne muscular dystrophy with a nonsense mutation in the dystrophin gene.*, 2023a. Available online: https://www.nice.org.uk/guidance/hst22 [Accessed.

National Institute for Health and Care Excellence. (2023b) *Onasemnogene abeparvovec for treating spinal muscular atrophy*, 2023b. Available online: [Accessed.

National Institute for Health and Care Excellence. (2025) *Lecanemab for treating mild cognitive impairment or mild dementia caused by Alzheimer's disease*, 2025. Available online: https://www.nice.org.uk/guidance/indevelopment/gid-ta11220 [Accessed.

Neumann, P. J., Sandberg Ea Fau - Araki, S. S., Araki Ss Fau - Kuntz, K. M., Kuntz Km Fau - Feeny, D., Feeny D Fau - Weinstein, M. C. & Weinstein, M. C. (2000) A comparison of HUI2 and HUI3 utility scores in Alzheimer's disease(0272-989X (Print)).

NHS Business Services Authority. (2025) *Programme budget categories (PBC).* 2025. Available online: https://www.nhsbsa.nhs.uk/access-our-data-products/epact2/report-information/programme-budget-categories-pbc [Accessed.

Papaioannou, D., Brazier, J. & Paisley, S. (2013) Systematic searching and selection of health state utility values from the literature. *Value Health*, 16(4), 686-95.

Park, M. H., Smith, S. C., Hendriks, A. A. J. & Black, N. (2019) Caregiver burden and quality of life 2 years after attendance at a memory clinic. *Int J Geriatr Psychiatry*, 34(5), 647-656.

Pearlin, L. I. (2010) The life course and the stress process: some conceptual comparisons. *J Gerontol B Psychol Sci Soc Sci*, 65B(2), 207-15.

Pennington, B. & Al-Janabi, H. (2024) Modelling Informal Carers' Health-Related Quality of Life: Challenges for Economic Evaluation. *Appl Health Econ Health Policy*, 22(1), 9-16.

Pennington, B., Davis, S. & Cranmer, H. (2026) Caring for and caring about in economic evaluation: modelling the family and caregiving effects., [Lecture].unpublished.

Pennington, B., Eaton, J., Hatswell, A. J. & Taylor, H. (2022) Carers' Health-Related Quality of Life in Global Health Technology Assessment: Guidance, Case Studies and Recommendations. *Pharmacoeconomics*, 40(9), 837-850.

Pennington, B., Hernández Alava, M. & Strong, M. (2025a) How Does Bereavement Affect the Health-Related Quality of Life of Household Members Who Do and Do Not Provide Unpaid Care? Difference-in-Difference Analyses Using the UK Household Longitudinal Survey. *Pharmacoeconomics*. 43(4), 389-402.

Pennington, B. M. (2020) Inclusion of Carer Health-Related Quality of Life in National Institute for Health and Care Excellence Appraisals. *Value Health*, 23(10), 1349-1357.

Pennington, B. M., Hernandez Alava, M. & Strong, M. (2025b) Unpaid Caring and Health-Related Quality of Life: Longitudinal Analysis of Understanding Society (the UK Household Longitudinal Survey). *Value Health*, 28(1), 138-147.

Potter, C. M., Peters, M., Cundell, M., McShane, R. & Fitzpatrick, R. (2023) Living well while providing support: validation of LTCQ-Carer for assessing informal carers' quality of life. *Qual Life Res*, 32(12), 3507-3520.

Roth, D. L., Fredman, L. & Haley, W. E. (2015) Informal caregiving and its impact on health: a reappraisal from population-based studies. *Gerontologist*, 55(2), 309-19.

Rummery, K. & Fine, M. (2012) Care: A Critical Review of Theory, Policy and Practice. *Social Policy & Administration*, 46(3), 321-343.

Schulz, R., Beach, S. R., Czaja, S. J., Martire, L. M. & Monin, J. K. (2020) Family Caregiving for Older Adults. *Annu Rev Psychol*, 71, 635-659.

Schulz, R. B., K.; Herbert, R.; (2008) Caregiving and bereavement., in Stroebe, M. H., RO.; Schut, H.; Stroebe. (ed), *Handbook of bereavement research and practice.* Washington DC.: American Psychological Association, 265-285.

Stöckel, J. & Bom, J. (2022) Revisiting longer-term health effects of informal caregiving: Evidence from the UK. *The Journal of the Economics of Ageing*, 21.

Sukchokpanich, P., Anusitviwat, C., Jarusriwanna, A., Kitcharanant, N. & Unnanuntana, A. (2023) Quality of Life and Depression Status of Caregivers of Patients with Femoral Neck or Intertrochanteric Femoral Fractures during the First Year after Fracture Treatment. *Orthop Surg*, 15(7), 1854-1861.

Sullivan, G. M. & Feinn, R. (2012) Using Effect Size—or Why the P Value Is Not Enough. *Journal of Graduate Medical Education*, 4(3), 279-282.

Tandvårds- och läkemedelsförmånsverket (2020) *Underlag för beslut i landstingen Vimizim (elosulfas alfa)*, 2020. Available online: https://www.tlv.se/lakemedel/kliniklakemedelsuppdraget/avslutade-halsoekonomiska-bedomningar/arkiv/2020-05-29-halsoekonomisk-bedomning-av-vimizim-vid-mps-iva-morquio-a-syndrom.html [Accessed.

The Dental and Pharmaceutical Benefits Agency (TLV). (2022) *Calculation and payment*, 2022. Available online: https://www.tlv.se/in-english/reports/arkiv/2023-01-27-calculation-and-payment.-continued-study-on-evaluation-methods-and--payment-models-for-new-medicines-such-as--atmps-and-precision-medicine----2022.html [Accessed.

Theakston, C. & Mott, D. J. (2025) Spillover Effects on Informal and Family Carers' Utility: A Systematic Literature Review, *HESG Winter 2025.* Brighton.

Torres, J. M., Mitchell, U. A., Sofrygin, O., Rudolph, K. E., Lopez-Ortega, M., Sharif, M. Z., Wong, R. & Glymour, M. M. (2021) Associations between spousal caregiving and health among older adults in Mexico: A targeted estimation approach. *Int J Geriatr Psychiatry*, 36(5), 775-783.

Tubeuf, S., Saloniki, E. C. & Cottrell, D. (2019) Parental Health Spillover in Cost-Effectiveness Analysis: Evidence from Self-Harming Adolescents in England. *Pharmacoeconomics*, 37(4), 513-530.

Ubel, P. A., Nord, E., Gold, M., Menzel, P., Prades, J.-L. P. & Richardson, J. (2000) Improving Value Measurement in Cost-Effectiveness Analysis. *Medical Care*, 38(9), 892-901.

University of Bristol. (2025) *ICECAP-A*, 2025. Available online: https://www.bristol.ac.uk/population-health-sciences/projects/icecap/icecap-a/ [Accessed.

van den Berg, B., Al, M., van Exel, J., Koopmanschap, M. & Brouwer, W. (2008) Economic valuation of informal care: conjoint analysis applied in a heterogeneous population of informal caregivers. *Value Health*, 11(7), 1041-50.

Walters, S. J. & Brazier, J. E. (2005) Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res*, 14(6), 1523-32.

Wittenberg, E., James, L. P. & Prosser, L. A. (2019) Spillover Effects on Caregivers' and Family Members' Utility: A Systematic Review of the Literature. *Pharmacoeconomics*, 37(4), 475-499.

Wittenberg, E. & Prosser, L. A. (2013) Disutility of illness for caregivers and families: a systematic review of the literature. *Pharmacoeconomics*, 31(6), 489-500.

Wittenberg, E., Ritter, G. A. & Prosser, L. A. (2013) Evidence of spillover of illness among household members: EQ-5D scores from a US sample. *Med Decis Making*, 33(2), 235-43.

Wrotek, A., Wrotek, O. & Jackowska, T. (2023) The Estimate of Parental Quality of Life Loss Due to Respiratory Syncytial Virus (RSV) Hospitalization. *Diseases*, 11(4).

Zhao, X., Zhang, Q., Ji, Y., Liu, H. & Lou, V. W. Q. (2023) Influence of spousal caregiving and living arrangement on depression among husband caregivers in rural China. *Aging Ment Health*, 27(7), 1266-1273.