

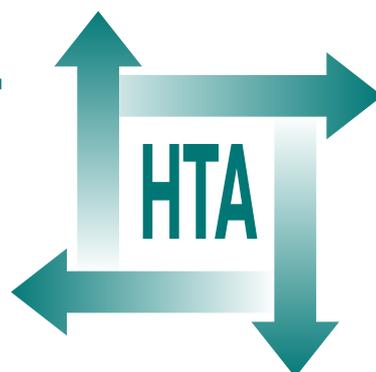
Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY Project

R Baker, I Bateman, C Donaldson,
M Jones-Lee, E Lancsar, G Loomes,
H Mason, M Odejar, JL Pinto Prades,
A Robinson, M Ryan, P Shackley,
R Smith, R Sugden and J Wildman
(the SVQ Research Team)



May 2010
10.3310/hta14270

Health Technology Assessment
NIHR HTA programme
www.hta.ac.uk





How to obtain copies of this and other HTA programme reports

An electronic version of this title, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (www.hta.ac.uk). A fully searchable DVD is also available (see below).

Printed copies of HTA journal series issues cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our despatch agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per issue and for the rest of the world £3 per issue.

How to order:

- fax (with **credit card details**)
- post (with **credit card details** or **cheque**)
- phone during office hours (**credit card** only).

Additionally the HTA website allows you to either print out your order or download a blank order form.

Contact details are as follows:

Synergie UK (HTA Department)
Digital House, The Loddon Centre
Wade Road
Basingstoke
Hants RG24 8QW

Email: orders@hta.ac.uk

Tel: 0845 812 4000 – ask for ‘HTA Payment Services’
(out-of-hours answer-phone service)

Fax: 0845 812 4001 – put ‘HTA Order’ on the fax header

Payment methods

Paying by cheque

If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *University of Southampton* and drawn on a bank with a UK address.

Paying by credit card

You can order using your credit card by phone, fax or post.

Subscriptions

NHS libraries can subscribe free of charge. Public libraries can subscribe at a reduced cost of £100 for each volume (normally comprising 40–50 titles). The commercial subscription rate is £400 per volume (addresses within the UK) and £600 per volume (addresses outside the UK). Please see our website for details. Subscriptions can be purchased only for the current or forthcoming volume.

How do I get a copy of HTA on DVD?

Please use the form on the HTA website (www.hta.ac.uk/htacd/index.shtml). *HTA on DVD* is currently free of charge worldwide.

The website also provides information about the HTA programme and lists the membership of the various committees.

Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY Project

R Baker,^{1,2} I Bateman,³ C Donaldson,^{1,2*}
M Jones-Lee,⁴ E Lancsar,^{1,4} G Loomes,⁵
H Mason,¹ M Odejar,⁶ JL Pinto Prades,^{7,8}
A Robinson,⁹ M Ryan,⁶ P Shackley,¹⁰
R Smith,¹¹ R Sugden⁵ and J Wildman⁴
(the SVQ Research Team)

¹Institute of Health and Society, Newcastle University,
Newcastle upon Tyne, UK

²Yunus Centre for Social Business and Health, Glasgow Caledonian
University, Glasgow, UK

³School of Environmental Sciences, University of East Anglia, Norwich, UK

⁴Newcastle University Business School, Newcastle upon Tyne, UK

⁵School of Economics, University of East Anglia, Norwich, UK

⁶Health Economics Research Unit, University of Aberdeen, Aberdeen, UK

⁷Department of Economics, University Pablo de Olavide, Sevilla, Spain

⁸Fundación Centro de Estudios Andaluces, Sevilla, Spain

⁹School of Medicine, Health Policy and Practice, University of East Anglia,
Norwich, UK

¹⁰School of Health and Related Research, University of Sheffield,
Sheffield, UK

¹¹Department of Public Health and Policy, London School of Hygiene &
Tropical Medicine, London, UK

*Corresponding author

Declared competing interests of authors: none

Published May 2010

DOI: 10.3310/hta14270

This report should be referenced as:

Baker R, Bateman I, Donaldson C, Jones-Lee M, Lancsar E, Loomes G, et al. Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY Project. *Health Technol Assess* 2010;**14**(27).

Health Technology Assessment is indexed and abstracted in *Index Medicus/MEDLINE*, *Excerpta Medica/EMBASE*, *Science Citation Index Expanded (SciSearch®)* and *Current Contents®/Clinical Medicine*.

NIHR Health Technology Assessment programme

The Health Technology Assessment (HTA) programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

Criteria for inclusion in the HTA journal series

Reports are published in the HTA journal series if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this issue of the journal was commissioned by the National Coordinating Centre for Research Methodology (NCCRM), and was formally transferred to the HTA programme in April 2007 under the newly established NIHR Methodology Panel. The HTA programme project number is 06/91/05. The contractual start date was in October 2004. The draft report began editorial review in March 2009 and was accepted for publication in June 2009. The commissioning brief was devised by the NCCRM who specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.

Editor-in-Chief: Professor Tom Walley CBE

Series Editors: Dr Martin Ashton-Key, Dr Aileen Clarke, Professor Chris Hyde,
Dr Tom Marshall, Dr John Powell, Dr Rob Riemsma and Professor Ken Stein

Editorial Contact: edit@southampton.ac.uk

ISSN 1366-5278

© 2010 Queen's Printer and Controller of HMSO

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (<http://www.publicationethics.org/>).

This journal may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to: NETSCC, Health Technology Assessment, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk), on behalf of NETSCC, HTA.

Printed on acid-free paper in the UK by Henry Ling Ltd, The Dorset Press, Dorchester



Abstract

Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY Project

R Baker,^{1,2} I Bateman,³ C Donaldson,^{1,2*} M Jones-Lee,⁴ E Lancsar,^{1,4}
G Loomes,⁵ H Mason,¹ M Odejar,⁶ JL Pinto Prades,^{7,8} A Robinson,⁹
M Ryan,⁶ P Shackley,¹⁰ R Smith,¹¹ R Sugden⁵ and J Wildman⁴
(the SVQ Research Team)

¹Institute of Health and Society, Newcastle University, Newcastle upon Tyne, UK

²Yunus Centre, Glasgow Caledonian University, Glasgow, UK

³School of Environmental Sciences, University of East Anglia, Norwich, UK

⁴Newcastle University Business School, Newcastle upon Tyne, UK

⁵School of Economics, University of East Anglia, Norwich, UK

⁶Health Economics Research Unit, University of Aberdeen, Aberdeen, UK

⁷Department of Economics, University Pablo de Olavide, Sevilla, Spain

⁸Fundación Centro de Estudios Andaluces, Sevilla, Spain

⁹School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, UK

¹⁰School of Health and Related Research, University of Sheffield, Sheffield, UK

¹¹Department of Public Health and Policy, London School of Hygiene & Tropical Medicine, London, UK

*Corresponding author

Objectives: To identify characteristics of beneficiaries of health care over which relative weights should be derived and to estimate relative weights to be attached to health gains according to characteristics of recipients of these gains (relativities study); and to assess the feasibility of estimating a willingness-to-pay (WTP)-based value of a quality-adjusted life-year (QALY) (valuation study).

Design: Two interview-based surveys were administered – one (for the relativities study) to a nationally representative sample of the population in England and the other (for the valuation study) to a smaller convenience sample.

Setting: The two surveys were administered by the National Centre for Social Research (NatCen) in respondents' homes.

Participants: 587 members of the public were interviewed for the relativities study and 409 for the valuation study.

Methods: In the relativities study, in-depth qualitative work and considerations of policy relevance resulted in the identification of age and severity of illness as relevant characteristics. Scenarios reflecting these, along with additional components reflecting gains in

QALYs, were presented to respondents in a series of pairwise choices using two types of question: discrete choice and matching. These questions were part of a longer questionnaire (including attitudinal and sociodemographic questions), which was administered face to face using a computer-assisted personal interview. In the valuation study, respondents were asked about their WTP to avoid/prevent different durations of headache or stomach illness and to value these states on a scale (death = 0; full health = 1) using standard gamble (SG) questions.

Results: Discrete choice results showed that age and severity variables did not have a strong impact on respondents' choices over and above the health (QALY) gains presented. In contrast, matching showed age and severity impacts to be strong: depending on method of aggregation, gains to some groups were weighted three to four times more highly than gains to others. Results from the WTP and SG questions were combined in different ways to arrive at values of a QALY. These vary from values which are in the vicinity of the current National Institute for Health and Clinical Excellence (NICE) threshold to extremely high values.

Conclusions: With respect to relative weights, more research is required to explore methodological differences with respect to age and severity weighting. On valuation, there are particular issues concerning the extent to which 'noise' and 'error' in people's

responses might generate extreme and unreliable figures. Methods of aggregation and measures of central tendency were issues in both weighting and valuation procedures and require further exploration.



Contents

List of abbreviations	vii	Using the willingness-to-pay and standard gamble data to derive a monetary value of a QALY	58
Executive summary	ix	Concluding remarks	62
1 Introduction	1	7 Conclusions	65
2 Weighting and valuing QALYs: literature and challenges	3	Weighting QALYs	65
Estimating relative weights for QALYs	3	The feasibility of estimating a monetary value of a QALY	68
Valuing QALYs in monetary terms	3	Implications for research and practice	68
3 Identification and presentation of attributes for QALY weights	7	Acknowledgements	71
Introduction	7	References	73
Methods	7	Appendix 1 Focus group introduction	77
Qualitative findings based on open- ended discussion	8	Appendix 2 Factor loadings	79
Results based on simple ranking	10	Appendix 3 Q sort statements and factor scores	81
Q methodology	11	Appendix 4 Diagrammatic approach to explaining QALYs	85
Summary	13	Appendix 5 Constraints on DCE design	87
Diagrammatic questions	14	Appendix 6 Version 2 of the willingness- to-pay/standard gamble questionnaire	89
Survey sample and administration	15	Appendix 7 Show cards for ranking	123
4 Discrete choice study	19	Appendix 8 Standard gamble cards	125
Basic design	19	Appendix 9 Differences between discrete choice and matching	131
Functional form and empirical approach ..	20	Health Technology Assessment reports published to date	135
Discrete choice results	21	Health Technology Assessment programme	157
Weighting QALYs using discrete choice data	24		
Further investigation of severity	32		
Concluding remarks	33		
5 Matching study	35		
Basic approach	35		
Methods	35		
Results	39		
Concluding remarks	45		
6 Exploring the feasibility of eliciting a monetary value of a QALY	47		
Basic strategy	47		
Overview of the questionnaire	47		
Implementation and results	48		





List of abbreviations

<i>AD</i>	age at death without treatment	NICE	National Institute for Health and Clinical Excellence
<i>AO</i>	age at onset	<i>QALY</i>	quality-adjusted life-year
CAPI	computer-assisted personal interview	<i>QALY</i>	QALYs gained from treatment
CV	compensating variation	<i>QL</i>	quality of life lost without treatment
DCE	discrete choice experiment	QoL	quality of life
DEFRA	Department for Environment, Food and Rural Affairs	SG	standard gamble
DfT	Department for Transport	SVQ	Social Value of a QALY
EQ-5D	EuroQol 5 dimensions	VAS	visual analogue scale
HSE	Health and Safety Executive	VPF	value of preventing a (statistical) fatality
N/A	not applicable	WTP	willingness to pay
NatCen	National Centre for Social Research		

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.



Executive summary

The National Institute for Clinical Excellence (NICE) was created in 1999 and renamed the National Institute for Health and Clinical Excellence (but still NICE) in 2005. One of NICE's major roles is to advise the rest of the National Health Service (NHS) in England about the adoption and retention of health interventions. In assessing health technologies, the NICE process involves the use of economic evaluations. These evaluations reflect the dominant paradigm in health economics, with new technologies assessed, usually relative to current practice, in terms of additional cost per quality-adjusted life-year (QALY) gained. Since the inception of NICE, therefore, and even in the wider health economics literature prior to 1999, this has raised the questions of whether QALYs gained by different beneficiaries of health care should be weighted equally and whether it is feasible to estimate a monetary value of a QALY.

The Social Value of a QALY (SVQ) Team based at Newcastle University and the Universities of East Anglia and Aberdeen were contracted from October 2004 to September 2007 to undertake two studies, each based on a survey of the population in England.

The first study was the 'relativities study', which aimed to identify characteristics of beneficiaries of health care over which relative weights are to be derived, and to estimate the relative weights to be attached to health gains according to the characteristics of the recipients of these gains.

The second study was the 'valuation study', which aimed to assess the feasibility of estimating a willingness-to-pay (WTP)-based value of a QALY.

Relativities study

To identify characteristics of beneficiaries, we took a predominantly qualitative approach, supplementing conventional qualitative methods with other techniques. This exploratory part of the project was an iterative process involving several phases and a range of methods including:

open-ended discussion in focus groups; simple ranking procedures; experimentation with sample questions; and a more complex ranking task, called Q methodology, involving card sorting. Qualitative findings were interpreted alongside the results of the other methods used.

This exploratory work, which lasted for over a year, taken together with considerations of policy relevance in consultation with representatives of NICE, resulted in the selection of the following attributes: age (both at onset of illness and at death) and severity of illness (with and without treatment). These were presented to respondents along with additional information representing gains in QALYs. They were put forward in the form of two types of question: discrete choice and matching.

Methods for presenting information about age, severity and health gain were devised in focus groups and tested in cognitive interviews. The most successful technique was achieved through the use of diagrams, which were first explained using an innovative, animated Microsoft POWERPOINT[®] presentation. Respondents were then presented with pairs of scenarios in a series of choices, examples of which are contained in the main body of this report.

The discrete choice and matching questions were part of a longer questionnaire (including attitudinal and sociodemographic questions), which was administered face to face using a computer-assisted personal interview. The survey was administered to a nationally-representative sample ($n = 587$) of the population in England by the National Centre for Social Research (NatGen) from February to April 2007.

Two innovative approaches were then applied to the derivation of weights:

- A discrete choice approach, whereby, having presented respondents with a series of pairwise choices, each comprising combinations of age, severity levels and components of the QALY, a function for the relative importance of these

attributes was estimated, this function being used as the basis of two approaches to deriving weights for age and severity.

- The 'QALY grid' approach, based on 'matching' (or 'person trade-off' questions), presenting respondents with a series of iterative pairwise choices over scenarios described in terms of age, severity levels and QALY components, where either age or severity was allowed to vary within a choice set. The scenarios in each set were initially described as having equal numbers of beneficiaries. Once an initial choice was made, the number of beneficiaries in the preferred option was reduced until the respondent was indifferent between alternatives presented. A set of questions was devised to allow derivation of relative weights for 20 different areas of the 'grid', covering five ranges of quality of life (from death to full health) and four age groups (between birth and 80 years of age).

The results of the discrete choice study show that age and severity variables did not have a strong impact on respondents' choices over and above the health (QALY) gains presented. In contrast, the QALY grid analysis of the matching data showed clear evidence of both age and severity impacts, although measures of these varied with the aggregation method: the more conservative method of aggregation showed gains to some groups being weighted up to 2.75 times more highly than gains to others, whereas an alternative method gave ratios that could differ by up to 4:1. Still, the patterns in the matching data are robust to the choice of method: there is a general tendency to give greater weight to younger people and those in poorer health, although somewhat less weight is given to the very youngest and those in the very poorest health with limited prospects of improvement.

Valuation study

The valuation study was a feasibility study conducted on a smaller convenience sample ($n = 409$). Respondents were asked about their WTP to avoid/prevent different durations of head pain or stomach illness, and to value these states on a 0–1 scale (death = 0; full health = 1) using a set of standard gamble (SG) questions.

Results from the WTP and SG questions were combined in different ways to arrive at values of a QALY. Depending on the aggregation procedure,

these vary from values which are in the vicinity of the current NICE threshold to extremely high values.

Conclusions

Implications for practice are limited because of the methodological nature of the research. The two main recommendations are as follows:

- On relativities, it could be said that it would be premature to propose any particular set of QALY weights at this point in time: before that point is reached, there is scope for both further reconciliation and replication. However, it might equally be argued that there is no scope for reconciliation and that we need to choose between the results in light of the caveats of the matching and discrete choice methods used.
- On valuation, it was never the intention to conduct a representative survey using a definitive method. The main recommendation, therefore, is that any future national sample survey should be preceded by further extensive qualitative research and cognitive testing to resolve the main questions identified in the present study.

The research recommendations arising from the study are as follows:

- The findings from the relativities study indicate that more work is required in the short term to reconcile the results obtained, although fundamental differences between the methods and results reported may challenge such reconciliation.
- In the longer term, and still with respect to relativities, further methodological research should attempt to account for some of the deficiencies of the methods (especially the particular discrete choice approach used in SVQ).
- Building on the results of the innovative methods that have been devised in this study to derive relative weights, further replication of these results is required to address this important policy issue.
- With respect to valuation, shorter-term work is required around the issues of aggregation, combining WTP and SG values and the appropriateness of different measures of central tendency.
- In the longer term, more qualitative and cognitive research is required around two issues

in particular: first, the problem of identifying health states to present to respondents which are 'minor enough' for people to be able to express their willingness to pay, but not so minor that respondents will accept only

minuscule risks of death when responding to SG type questions; and second, but related to the first, the extent to which 'noise' and 'error' in people's responses might generate extreme and unreliable figures.

Chapter I

Introduction

Over the past 25 years, the quality-adjusted life-year (QALY) has become the dominant measure of benefit assessment in health economic evaluation.^{1,2} Its use is now widespread, particularly in the various health technology assessment agencies around the globe, and most notably in the UK through the assessment procedures undertaken by the National Institute for Health and Clinical Excellence (NICE).³ Nevertheless, almost since the QALY was introduced to the health economics literature, the importance of the context in which health gains are produced has been discussed, raising the question of whether ‘a QALY is a QALY’.^{4,5} Williams has said: ‘there is nothing in the QALY approach that requires QALYs to be used in a maximising context ... more complex rules will almost certainly be needed if collective priority-setting is to reflect the views of the general public.’⁶ This issue has come to the fore again recently through NICE. A prescribed task of NICE is to assess health interventions in terms of their health gains relative to their costs, and to make recommendations as to whether or not an intervention should be adopted by the rest of the National Health Service (NHS) in England.

However, it has been recognised that the Appraisal Committee at NICE will take characteristics of beneficiaries of such interventions into account in its deliberations.⁷ This raises an important policy question concerning whether quantitative estimates, reflecting the relative weight to be attached to health gains derived by different beneficiaries, can be elicited from a survey of the general public and thus be used to assist the NICE process.

Furthermore, the NICE approach involves making recommendations based on evaluations of single health-care interventions, which inevitably involve judgements about whether the QALYs gained are worthwhile.^{8,9} If it is thought necessary to have such benefits and costs expressed in a common metric, usually money, this raises a second policy question: what is the monetary value of a QALY? The threshold monetary value of a QALY used by NICE was queried by the House of Commons Health Select Committee in 2007,¹⁰ thus highlighting the importance of this policy issue.

The Social Value of a QALY (SVQ) Team was contracted from October 2004 to January 2008 to undertake two studies, each based on a survey of the population in England. The first was the ‘relativities study’, which had the following aims:

- to identify characteristics of beneficiaries of health care over which relative weights were to be derived; and
- to estimate the relative weights to be attached to health gains according to the characteristics of the recipients of these gains.

The second was the ‘valuation study’, which had the following aim:

- to assess the feasibility of deriving a monetary value of a QALY.

The rest of this report is organised into six further chapters as follows. We begin with a brief review of the literature on QALY weights and survey-based approaches to assessing the monetary value of a QALY, highlighting articles of particular interest and concluding with the challenges encountered in this project. Chapters 3–5 focus on the relativities study. In Chapter 3 we describe the methods and results of exploratory and developmental work to identify attributes and the development of a diagrammatic approach to the presentation of survey questions. We adopted two methodological approaches (which nevertheless share some commonalities) to the elicitation of preference data to be used in estimating relative weights, with respondents to the relativities survey answering both matching (or person trade-off) and discrete choice questions. The design, analysis and results of the discrete choice part of the survey are described in Chapter 4, with the same aspects of the matching part of the survey being outlined in Chapter 5. The valuation survey is described in Chapter 6. Two further introductory remarks are worth making at this point. Despite the extensive analyses undertaken to date, the results are nevertheless preliminary; for example, there is considerable scope to link the two data sets from the relativities survey, which may help to resolve some of the issues raised later in the report. An initial attempt at this is presented in Chapter 7.

Until such analyses have been completed, it is fair to say that the groups within the Team [broadly, Newcastle-led and University of East Anglia (UEA)-led] have differing perspectives on the two main approaches to the relativities work; the Newcastle-led group thinks that both exercises have their

merits, while the UEA-led group stands by the methods and results from the matching study. The report, therefore, concludes with further details on these differing perspectives, as well as a brief discussion and recommendations, largely for future research as opposed to current policy.

Chapter 2

Weighting and valuing QALYs: literature and challenges

Estimating relative weights for QALYs

Several authors have discussed the theoretical, ethical and practical issues around distributional weights for QALYs,¹¹⁻¹³ and there have been a number of attempts to estimate weights.^{11,14,15}

Dolan *et al.*¹⁶ provide the most comprehensive review in this field to date. Using a 'citation pearl growing' search strategy, they identify 78 papers, dated 2001 or earlier, of which 64 include empirical data. There is growing evidence from (mainly survey-based) studies of the general population that the number of QALYs gained is likely to be traded off against other factors. Efforts to identify these other factors have indicated a wide range of possibilities, but there remain inconsistencies and contradictory findings. The list of factors identified by Dolan *et al.* (drawn from both the empirical and theoretical literature) are as follows: age, severity of illness (or starting point health state), end point/final health state, culpability/responsibility for ill health, having dependants, socioeconomic characteristics, gender, ethnicity, inequalities in health, and the concentration or dispersion of the distribution of a fixed health gain. The reviewers conclude that despite a growing body of literature, there are contradictory findings, many studies involve small samples and few attempt to estimate weights. Earlier, Schwappach¹⁷ described two categories of factors that could influence social value: (1) characteristics of beneficiaries, and (2) characteristics of the intervention's effect on patients' health, adding the factors of prior health consumption of patients, the duration of benefit and whether the gain in health is an improvement or the prevention of a decline.

In the period since these reviews, there have been further contributions to the literature. Two of these used discrete choice data, examining attributes such as age, culpability (e.g. related to alcohol consumption), expected length of survival, time on waiting list and whether a previous transplant had been received;¹⁸ and lifestyle, socioeconomic status,

age, life expectancy, quality of life (QoL) after treatment and level of past use of health care.¹⁹ The results indicate that several factors, in addition to health gain, influence people's choices. However, unlike the studies described in this report, these studies were either condition specific or not based on a population sample of respondents.

Projects attempting to derive QALY weights are faced with three significant challenges: identifying characteristics of beneficiaries over which weights *should* be derived; designing and presenting questions so that respondents can understand complexities and make choices; and elicitation of quantitative preference data from members of the general public to allow the estimation of QALY weights.

Designing questionnaires that respondents can engage with was a particular challenge in this study because our aim was to estimate the relative value of different types of decontextualised, generic QALY gains. Without context, however, questions can seem overly abstract to respondents. There is also evidence that different 'types' of decontextualised QALY (e.g. life-saving or QoL-enhancing QALYs) will be regarded differently,^{20,21} and so the presentation of questions needed to be flexible enough to incorporate different QALY types.

Valuing QALYs in monetary terms

The concept of willingness to pay (WTP) has existed for a long time.^{22,23} However, not until the 1980s did government Transport Departments worldwide consider using the method to value lives saved from safety projects, rather than the gross output ('productivity') approach used previously.²⁴ Arguably, the most natural measure of the extent of a person's preference for anything is the maximum amount that they would be willing to pay for it. Under what has naturally come to be known as the 'willingness-to-pay' approach to valuation of safety,

one seeks to establish the maximum amounts that those affected would individually be willing to pay for (typically small) improvements in their own and others' safety. These amounts are then aggregated across individuals to arrive at an overall value for the safety improvement concerned, thus reflecting society's overall resource constraint.

Estimating a WTP-based monetary value of a QALY can also be viewed as a group-aggregate WTP for marginal gains in health, at least in the case of a randomly-selected sample of the public. Indeed, this argument has been used in promoting an insurance-based approach to valuing publicly-provided health care, whereby respondents are informed of the probabilities of needing care, as well as of it being successful, before providing a valuation.^{25,26}

The WTP method was first applied in health to value heart attack prevention.²⁷ Subsequently, there were few such studies in health, probably as a result of the view that such monetary valuation was unethical. In addition, the use of WTP to inform decisions about allocation of health care, which is supposed to be on the basis of (some notion of) need, may look problematic because WTP is obviously associated with ability to pay. However, it has been shown that this need not impede the use of WTP in health economic evaluation²⁸ and that, indeed, QALYs suffer from the same phenomenon.²⁹ Since the early 1990s, the feasibility of using WTP in health economics has again been recognised,^{25,30} and more studies have been undertaken.^{31,32}

Thus, in health, WTP methods historically addressed decision-making dilemmas assessing relative utility of treatments at two main levels: (1) for a given group of patients (involving elicitation of values from samples of such patients), and (2) across disparate programmes funded by geographically-defined health organisations (involving elicitation from the community of WTP values for each programme at stake). Methods have been developed which work well in terms of WTP values reflecting patient preferences.³³ In the latter area, methods have been more problematic, but are improving.^{34,35} As in other public sector areas, results have been mixed on how sensitive WTP responses are to the size of the good (i.e. the health change/numbers treated) on offer to respondents³⁶⁻³⁸ and to other aspects related to 'framing' and programme information presented to respondents.^{39,40} However, innovations in valuing

safety improvements derived by Carthy *et al.*⁴¹ have shown promise in overcoming these issues. Methods based on these developments, consistent with the notion of starting with 'marginal' gains (in this case, in health), are described below. For valuing a QALY, the challenge is to start with a health gain or detriment which is not so large as to hit respondents' budget constraints, but not so small that respondents are unwilling to trade it off against a risky situation in the standard gamble (SG) type question, also involved in a procedure of the sort devised by Carthy *et al.*⁴¹

Through the 1990s, development of national-level technology assessment agencies led to calls for monetary values of a QALY to aid decision making at a national level.^{42,43} In the UK, there has been significant debate about the empirical basis of the cost-per-QALY threshold above/below which NICE would recommend rejection/adoption of a therapy by the NHS. For example, the proceedings of the 2007 House of Commons Health Select Committee criticise the current NICE threshold on the bases that it '... is not based on empirical research and is not directly related to the NHS budget, nor is it at the same level as that used by PCTs [primary care trusts] in providing treatments not assessed by NICE.' Following previous literature, the implication is that, once a budget is set for the NHS (such budget setting not being the responsibility of NICE), we can infer a threshold by observing the cost per QALY of treatments which are funded vis-à-vis those that are not.⁴⁴

Two related responses to these arguments can be made: (1) it is well known that the NHS at the local level is not systematic in how it makes such decisions, at least in economics terms;⁴⁵ and (2) because the NHS is not good at curtailing existing therapies which are poor value for money, it is not really known whether the marginal cost per QALY within the rest of the NHS is indeed out of line with (i.e. lower than) that used by NICE.

Given these significant challenges to 'discovering' a threshold, an alternative is to ask members of the public about their WTP for health gains. It may be thought that asking individuals about their WTP for such health gains from their own pockets would not be relevant to the issue of establishing a threshold value contingent on the size of the health-care budget. Indeed, Culyer *et al.*⁴⁴ (p. 57) state 'Therefore, information about how much an individual or society values improvements in health (i.e. their WTP for a QALY) is not at all relevant to

the NICE remit. These values could only be used as the appropriate threshold by NICE if it were also given responsibility to set the NHS budget.'

This is an internally consistent position. However, it does not diminish the importance of trying to establish what value(s) people actually do place on QALY gains. There are two main reasons for this.

First, in a democratic society there is a case to be made for ensuring that the government's budget-setting process should, as far as possible, be informed by the preferences of members of the public. While it is reasonable to expect negotiations between the Treasury and Department of Health to take account of a number of factors, information regarding the public's WTP for health care should arguably constitute an important consideration.

Second, a theoretical argument is that, when assessing WTP questions in surveys, if respondents think of the NHS as being at full efficiency and unable to provide more services (or QALYs) without extra payments being made, then expressed WTP amounts would be a reasonable representation of a value of a QALY at the margin for the NHS and not far removed from what a budget-holder, like a PCT, might say is the value (if PCTs used QALYs

and if they behaved in an economically rational and QALY-maximising fashion!).

If the present study were to suggest that eliciting a robust monetary value of a QALY is feasible, and if a suitably representative sample survey were then undertaken, it would make a significant contribution to policy with respect to thresholds.

However, suitable existing evidence is scant and of variable quality. Some estimates have been made of the value of a QALY based either on modelling approaches or on survey research.^{46,47} Modelling studies have been reviewed elsewhere and values of a QALY vary greatly depending on how the data are manipulated.⁴⁸ Moreover, survey work on the value of a QALY has been limited. Typically, individuals have been asked about their WTP for health gains for which quality adjustment factors have been obtained from another sample without fully adjusting for uncertainty (i.e. by presenting scenarios involving certain gains in QoL) and, in some cases, eliciting values from patients and not from members of the general public.^{47,49,50} Only one such estimate exists for a European country.⁴⁷ The research undertaken in this study, therefore, represents a significant advance in the methods in this area.

Chapter 3

Identification and presentation of attributes for QALY weights

Introduction

The first objective of the relativities study, 'to identify characteristics of beneficiaries of health care over which relative weights are to be derived', requires qualitative enquiry. This precedes the quantitative estimation of the relative importance placed on those characteristics, once established. While the quantitative study is reported in two subsequent sections, reflecting two methodological approaches, these approaches share a common empirical foundation based on in-depth exploratory and developmental research.

The selection of attributes for inclusion in the relativities study is crucial. The inclusion of an attribute without a robust rationale for doing so, or the omission of an important attribute, will lead to misleading conclusions. Qualitative techniques are increasingly used to establish appropriate attributes, particularly for discrete choice studies.⁵¹ In this study we have taken a predominantly qualitative approach and supplemented conventional methods with other techniques. This exploratory phase of the project was an iterative process, involving three waves of focus groups with members of the general public, and use of a range of methods including: open-ended discussion; simple ranking procedures; experimentation with sample questions; and a more complex ranking task involving card sorting (Q methodology). Qualitative findings were interpreted alongside the results of the other methods used.

Most of this chapter is taken up by reporting on the three main methods used to identify the most important characteristics of beneficiaries, followed by a description of the development of methods used to present information to respondents using innovative diagrammatic representations, before, finally, describing the format of the questionnaires, presented using a computer-assisted personal interview (CAPI).

Methods

Focus groups were facilitated by two or more members of the research team depending on the size of the group. Discussions were introduced and guided by the focus group leader. Other researchers were available to help distribute materials and answer questions during individual exercises. Group discussions were recorded using a digital voice recorder and transcribed verbatim. All focus groups received an introductory description of the project and the problem at hand. This took the form of a brief presentation followed by the opportunity to ask questions. Participants were then guided through two or more tasks.

During the first wave of focus groups, we adopted open-ended qualitative techniques to elicit views and to probe responses. Respondents were advised that resources to provide health services are constrained and, as such, difficult choices must be made about the types of treatments and interventions that are provided by the NHS and, by implication, those that are not. The focus group leader introduced notions of scarcity of resources, carefully, and in simple terms (see Appendix 1). Essentially respondents were asked to accept the inevitability of rationing (although that term was not used) and that with or without the views of the general public, priority setting will happen. There was positive acceptance of these facts and respondents were comfortable proceeding on that basis. They were asked to suggest what sorts of things should be taken into account when such decisions are made.

Despite interesting discussions (a summary of which is presented below), participants often had difficulty absorbing and expressing opinions about the concepts we wanted them to explore. Generally, respondents readily proposed issues such as the size of the health gain, cost, QoL and life extension, but when asked to 'go beyond'

these concepts there was difficulty (or perhaps reluctance). If the group discussion then stalled, examples of possible issues for discussion were suggested by group facilitators, but the results of this approach raised concerns about leading respondents and endowing particular attributes with validity simply by mentioning them.

In the second wave of focus groups, we included a simple ranking task both as a means of generating some crude data and to stimulate discussion. Respondents were asked to rank order a set of 10 cards printed with issues that might be considered in priority setting, such as 'quality of life of patients before treatment' or 'the social class of patients typically affected'. Participants in these focus groups were also presented with some examples of the types of questions that would be used in the quantitative study, stimulating debate about the inclusion of different attributes as well as providing valuable information on the appropriateness of different modes of presentation.

In the third wave of focus groups, we introduced Q sort techniques which are sufficiently distinct to warrant a separate section and this follows the general findings. Q sorting involves arranging a number of cards, printed with statements about the topic, according to an instruction such as 'from most agree to most disagree'. Allowing focus groups to begin with individual Q sort activities and following this with discussion enabled respondents to express their views (via the Q sort) before engaging in discussion with others. They also entail a common stimulus set (in this case 46 cards printed with different statements about the issue at hand). These results are, therefore, unencumbered by the input of the focus group leader, or by the views of others during discussion.

Focus groups were conducted with groups of between 4 and 10 respondents and organised in three waves of data collection in Newcastle upon Tyne and Norwich between March and November 2005. Participants in the Newcastle groups were recruited through a social research company (NWA Social and Market Research) based in the north-east of England and £20 was paid to each participant in recognition of their travel expenses and time. Participants in Norwich were recruited from an existing university database of members of the general public who had consented to be contacted for research. A total of 126 respondents (42 in Norwich and 84 in Newcastle) took part in focus groups.

Qualitative findings based on open-ended discussion

As already stated, the use of open-ended qualitative techniques was only a partial success. This may reflect the fact that respondents do not necessarily have a readily articulated account of their views on such complex issues, or that they were unsure about the kinds of things we were asking them to consider. In general, respondents were far more comfortable talking about health-related characteristics than they were discussing social or personal characteristics of the beneficiaries of health care. We do not report a complete, formal qualitative analysis because of the nature of these data, our own objectives and word limits. Instead, a summary is given of the nature of the discussion on each issue, including some brief illustrative quotes, and the results of the simple ranking exercise are included.

Age

Age was important to participants for a variety of reasons (which are also well documented in the literature). In different accounts, the young were favoured because of: their longer life expectancy; fair innings arguments; current and future contributions to society; and productivity. This was not uncontested; others argued the deservedness of older people, whose life-long contributions to the financing of the health services should be recognised:

I would still veer towards [the] 7 year old because the 7 year old's got all its life in front of it, whereas the 70 year old has had 70 years of life.

Focus group respondent, Newcastle,
May 2005

Age is also a 'proxy' for a range of other characteristics. Whether or not potential patients are economically active has already been mentioned, but patients' social and familial networks were also linked to their age. The most obvious connection made was ages when patients are likely to be in their childbearing and childrearing years.

'The average age of patients at time of illness' was ranked 4th highest of 10 in the simple ranking exercise (below cards listing QoL, life expectancy and whether or not other treatments were available).

Dependants

In early focus groups, discussion about prioritising health care for people with dependants sparked off quite significant disagreements. As well as the view that we had expected, i.e. that some respondents might attach positive weight to health gains to people with young children or other dependants, there was also strong opposition to this view which rejected the diversion of health care funds in favour of people who had made a choice to have children:

... first of all I thought that might influence my decision if they had dependants ... but then I thought about it in another way, it seems unfair if you do have dependants that just because you've got dependants its influencing your decisions ... that you get that intervention, so it seems unfair.

Focus group respondent, Newcastle,
May 2005

Such positive weighting of health gain for people with dependants was also seen as discriminatory against the childless. Others pointed out that people can be good and bad 'carers' but that favouring those with dependants seems to indicate a moral worthiness to this role in exclusion of other roles in a community:

Making moral judgements is dodgy ... part of it seems to be about how worthy somebody is to be given resources. Because they may have dependants but their quality of care might not be great to those dependants, not all parents are good at parenting. Whereas their value in life might be that they're very good at their job, or they're very good as a friend.

Focus group respondent, Newcastle,
May 2005

In the simple ranking, the average rank for 'Whether or not the patients have dependants' was 7/10.

Lifestyle

In a similar vein to the comments on dependants, the issue of whether lifestyle (or 'culpability') should be taken into account divided respondents. The process of discussion and the views of others also seemed to affect respondents' stated views. Here, issues of choice and control, addiction and social/environmental influences were mentioned. Respondents appreciated the difficulty many addicts have in quitting, as well as the fact that,

for older people, information about the risks of some behaviours had not been available (in the following quotes, different focus group respondents are distinguished using letters, e.g. respondent G, respondent B):

- G Lifestyle is a much more complex thing than just yes or no choice.
- B Yes, its what people, people like smokers they're ...
- G Stressed or ...
- B ... they're stressed or living in absolute poverty and have to ... you know, that's why they've got ...
- K And actually smoking is very difficult to give up. It's no good saying, oh well people smoke therefore it's not right, it's very difficult to give up.

Focus group discussion, Newcastle,
May 2005

A small number wanted to prioritise those who take care of their own health above those who smoke or drink, but most had difficulty sustaining a logical argument in the face of disagreement. Obesity seemed to generate different views than smoking and alcohol and was used as an example of the 'slippery slope' down which such discussion can descend. The culpability argument was applied to a wide range of activities, including sports injuries, for example:

- P But it's starting on the slippery slope isn't it? Where do you draw the line?
- N It doesn't harm them drinking a little bit.
- P Going on from alcohol to diet ... because people are fat should we penalise them?

Focus group discussion, Newcastle,
May 2005

Discussion about liver replacement (which usually centred on the 'George Best case') resulted in more respondents wanting to incorporate lifestyle into decisions, but the dominant view was that everyone should be entitled to a first chance at treatment, regardless of past lifestyle. After that, failing to follow medical advice would be regarded negatively. Discussions usually concluded (not necessarily with consensus) that this is a problematic area and that prevention and health education are important areas for funding.

In the simple ranking task, 'Whether or not the patients live a healthy lifestyle' was ranked in the middle, at 5th on average.

Socioeconomic status: some conflicting evidence

Findings from open discussions about the importance of socioeconomic status were often difficult to interpret, in particular because a range of issues are conflated, although the simple ranking and Q sort data would suggest that socioeconomic group should not be included as an attribute.

There are several socioeconomic issues which were mentioned. The first (and probably the main one we had anticipated) is linked to alleviation of deprivation and the prioritisation of interventions aimed at this over other interventions which may have a higher potential health gain, but which do not deal with inequalities. Individuals' ability to pay was commonly mentioned in this area and appeared to cloud the issue of inequality – the rich being able to pay being seen as a pragmatic solution rather than an issue of equity. Some respondents were adamant that socioeconomic considerations should not be taken into account, and cited the foundations of the NHS and equal treatment of all. In group discussions, the different socioeconomic issues were generally not delineated or articulated clearly. In one group, a respondent (who was also a health professional) raised the issue of inequalities. Otherwise people argued for dealing with poverty, not for giving poor people 'preferential treatment'. Several respondents were appalled at the suggestion that health care might be prejudiced against people with higher socioeconomic status, especially as they had contributed to the NHS through higher taxes.

'The social class of patients typically affected' was ranked last (10/10) on average in the simple ranking. There were, however, some concerns that social class, without further explanation, was being interpreted by some to mean discriminating *against* those in more deprived groups. (The Q sort statements made more explicit the 'direction of effect'.)

Quality of life of beneficiaries

This health-related factor was discussed at some length in all focus groups. There were two main arguments. The first related to the relationship between length of life and QoL, the thrust of opinion focusing on the unnecessary extension of life in older people experiencing poor levels of QoL:

You wouldn't want to live longer in a worse health state, quality of life is the important thing.

Focus group participant, Norwich,
March 2005

The second argument related to the 'starting point' QoL before treatment and the relationship between that starting point and the amount of QoL gained through treatment. Here, some respondents observed that an improvement in QoL for people in very poor health would be more important than an identical improvement in QoL for people in relatively good health.

The following illustrative quote refers to an example of a question in which QoL is represented on a scale from 0 to 100, using percentages for ease:

I went for (choice) 'A' because I thought that a jump from 20% to 40% would make a huge difference, a bigger difference than from 70% to 90%. I can imagine 70% being a healthy state that you could quite easily live and not have to take too many treatments and that kind of thing, whereas 20% is pretty close to death.

Focus group participant, Newcastle,
May 2005

Results based on simple ranking

A subgroup of 19 respondents (aged 20–62, 10 female) rank ordered a set of possible attributes according to their importance and discussed their rankings. They were also invited to add any additional attributes (writing them onto blank cards provided) and incorporate those into their ranking. *Table 1* presents the average ranking of each item. This is only illustrative; respondents' rankings are not intended to be interpreted in isolation of their comments and the results of other methods.

Respondents' comments during this task revealed multiple understandings of the attributes as well as a small number of common views. Respondents were comfortable and confident talking about health-related attributes and less so when discussing non-health-related attributes. Socioeconomic status and gender were considered irrelevant to issues of prioritisation by all respondents. Whether or not patients had had a lot of health care in the past was often construed

TABLE 1 Simple ranking exercise

Rank	Average rank ^a	Attributes
1	2.9	QoL of patients before treatment
2	3.1	Whether there is no other treatment available
3	3.2	The life expectancy of patients before treatment
4	4.4	The average age of patients at time of illness
5	5.9	Whether or not the patients live a healthy lifestyle
6	6.1	Whether or not the patients have had a lot of health care in the past
7	6.3	Whether or not the patients have dependants
8	7.4	Whether or not the patients are currently working
9	9.0	The gender of patients typically affected
10	9.2	The social class of patients typically affected

a Average rank is simply the mean ranking given to the listed attributes ($n = 19$).

as the health service having failed them and issues of 'orphan drugs' were difficult for respondents to appreciate and were not covered by the lack of any other treatment.

While the qualitative data and simple ranking data provide a good grounding in the issues in question, we felt it was insufficient for the selection of attributes and so incorporated a third method, Q methodology. This is relatively unfamiliar to most and so requires separate explanation below.

Q methodology

The basic features of Q methodology

Q methodology⁵²⁻⁵⁴ is used to study the nature of views, opinions and beliefs. It is a useful addition to qualitative methods, especially where respondents do not necessarily have readily articulated accounts of their views on a topic.

The two main features of a Q study are the data collection method – which is based on a card sorting technique (the 'Q sort') – and a form of factor analysis which is used to analyse patterns between the card sorts to reveal a small number of underlying perspectives. There are several key terms that are used in Q studies. The 'Q sort' provides the primary data source in Q methodology. Respondents sort a set of statement cards, known as the 'Q set'. The Q set comprises a number of statements which cover the range of viewpoints and opinions on the particular topic of interest. Respondents consider each card in turn and assign it in a quick, initial sort, to one of three piles: agree, disagree or neutral. A more detailed arrangement of cards then follows, using

a grid such as the one reproduced in *Figure 1*. Factors are the result of Q analysis, the aim of which is to identify shared views and meanings that exist around a topic (via correlations between the positioning of cards by respondents).

Each space in the grid indicates the positioning of a card on the continuum from -5 to $+5$. Two items are placed in the $+/-5$ positions, four items in the '4' and '3' positions and so on. Making use of the three initial piles, respondents are asked to consider the cards in their 'agree' pile, select two cards that they 'most agree' with and place these in the $+5$ column. Next, selecting from the cards that they disagree with, respondents are asked to select the two cards that they 'most disagree' with, and place those cards in the -5 column. This process is repeated until all cards are placed (46 in this example), finishing at the centre of the distribution. Often the Q data (i.e. the positioning of the cards in the Q sort) are supplemented by a brief interview. In this study the Q sorts were followed by group discussion.

Factors and factor loadings in Q methodology

In Q methodology, 'factors' are distinct accounts, each one a shared point of view, relating to the topic studied, based on the correlations between respondents' Q sorts.

There are several types of information of interest in the interpretation of factors. The main source of information is a 'collective' Q sort (known as a *factor array*) for each factor, which is calculated from the individual Q sorts making up that factor and based on weighted averages. In other words,

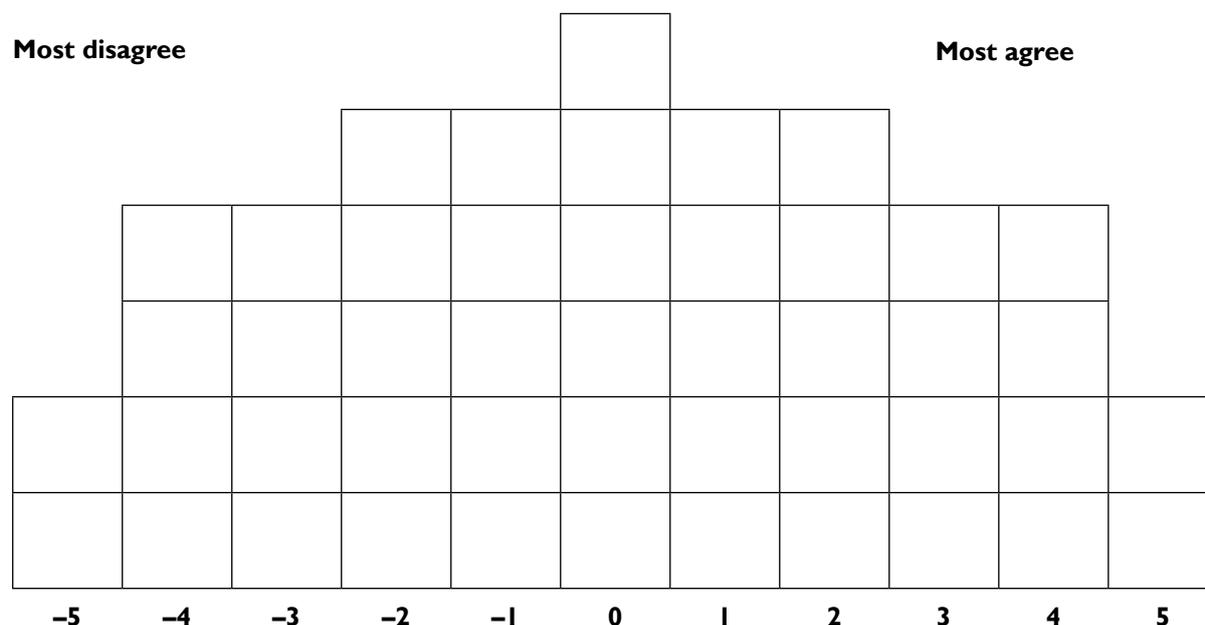


FIGURE 1 Q sort response grid.

for each factor, all of the statement cards can be placed on the Q grid, representing the relative importance of each statement to each factor. The statements placed in the +/-5 and +/-4 positions (representing strongly held views) are important in the interpretation of factors.

Information is also generated that identifies those statements which significantly distinguish factors, as well as consensus statements which are positioned similarly across factors. ‘Factor loadings’ give us further information about the correlation between each individual’s own Q sort and each factor – see Appendix 2, which presents factor loadings for each respondent. Qualitative data collected during and after the Q sorts, including spoken and written comments, are also used to help understand the meanings contained in the Q sorts.

The SVQ Q study

Statements for inclusion in the Q sort were taken from the first wave of focus groups, conducted in March 2005. Using the audio-recordings of the group discussions, opinions expressed by respondents which related to the topic of interest were listed. A small number of statements thought

to be of interest to the research questions, but not raised in the groups, were added. Duplicate statements were then deleted, selected statements were reworded to make them clearer, and the list was reduced to a set of 46 statements through a process of discussion within the research team. Pilot Q sorts were completed by a sample of the general public ($n = 26$) and a sample of Newcastle University staff members ($n = 23$). During this pilot, respondents were asked for comments about the set of statements; and in particular, to suggest anything that they felt was relevant that was not included in the statements. A revised set of statements was produced based on their comments. This final set of statements is reproduced in Appendix 3.

In the final wave of focus groups, a subgroup of 27 respondents (aged 20–84, 16 male) sorted a set of 46 cards printed with statements about the topic in question, according to their agreement with them.

SVQ Q findings

A three-factor solution was derived using centroid factor analysis and judgemental rotation.⁵⁵ Further information on the analysis is available from the authors.

Factor 1: egalitarianism

The first factor identified appears to be an egalitarian account, rejecting prioritisation of health care according to characteristics such as social class, lifestyle or whether respondents have dependants. It is an account which is concerned not primarily with outcome but with *entitlement*, and equal access to health care for everyone.

Table 2 lists the statements of most importance to factor 1. In the top half of the table are statements that found strong agreement in factor 1 (i.e. these were statements placed at +4 and +5 positions on the Q sort grid by respondents associated with this factor). The statements in the bottom half of the table (below the emboldened line) are those that provoked strong disagreement. All the statements of importance in this factor (with the possible exception of 6) reflect an egalitarian position: no distinction should be made between age groups, socioeconomic groups or those without dependants. In keeping the statements in the bottom half of the table, the -4 and -5 positions are rejected by factor 1, supporting the interpretation. This factor refuses prioritisation on the basis of the characteristics of beneficiaries even when, as in statement 14, there is some implied gain in overall health in doing so.

Factor 2: health benefits

Factor 2 reveals a somewhat different point of view. This second account puts emphasis on outcome and the size of the health benefits – as revealed through the placing of statements 29, 31, and 44 in the ‘most agree’ columns of the grid for this factor (Table 3). Interestingly, this is coupled with a rejection of any statements that make reference to socioeconomic and financial issues. All of the statements placed in the -4 and -5 positions (‘most disagree’) are of this nature. Preventive health care is also important in this, as it is in all three accounts.

Factor 3: children and experts

A third factor shared some views with those already described, but is distinguished by a concern for children’s health and a belief that health prioritisation decisions should be made by experts. Examination of the full set of statements and factor scores (see Appendix 3) reveals that factor 3 often shares views with factors 1 and 2 or occupies a space between them. However, focusing specifically on significantly distinguishing statements, we can detect the views that set this account apart – for example, statement 35 (‘The decisions about which

services to fund, and how to spend NHS money should be made by the experts’), placed in the +3 position for factor 3 and in -3 and 0 for factors 1 and 2 respectively), and statement 13 (Table 4), placed in the +4 position for factor 3 but 0 for both factors 1 and 2.

Once again, there is a notable reluctance to prioritise on the basis of other factors evidenced by the statements that are rejected.

Overall Q findings

Respondents associated with all three factors thought that health care should be based on some concept of need and not on other factors, such as socioeconomic characteristics (statement 30) or lifestyle factors (statement 41). There was a rejection of socioeconomic issues and statements about lifestyle, and dependants were placed in the middle (irrelevant) or at the ‘disagree’ end of the scale in all three factors.

Summary

Based on our qualitative and Q methodological enquiry, we rejected both lifestyle and socioeconomic status as attributes. Age, QoL and length of life were clearly issues of importance. The issue of whether beneficiaries of health care have dependants was slightly more difficult to resolve, but ultimately we took account of considerations of policy relevance, in consultation with representatives of NICE. It is difficult to conceive of a situation where NICE would recommend an intervention be made available only to people with dependants.

Arriving at a list of key attributes for inclusion in quantitative survey questions is a difficult process. Increasingly, qualitative methods are built into the early stages of study designs. In the context of this study, the use of open-ended qualitative methods alone proved insufficient to determine the attributes, perhaps because of the complexity of the subject matter. Respondents appeared to be led by the suggestions of the focus group facilitator, and there remained uncertainties around particular attributes (such as socioeconomic status) following analysis of the qualitative data. In this case, Q methodology provided additional structure and the opportunity to use a standardised stimulus in both the generation of data and the analysis. With respect to socioeconomic status, for example, we were able

TABLE 2 Salient statements for factor 1

Number	Statement	Factor score
11 ^b	Life is equally valuable whether you are young or old.	+5
15 ^b	Everybody, no matter what you are, whether you are young or old, should get the same access to and choice of treatment.	+5
41	Health care should be based on <i>need</i> , not on social circumstances, or addiction or weight or smoking or anything else.	+4
6 ^b	If someone is given treatment, like George Best, and then abuses their treatment, they should not be given repeated chances. If there are finite resources and a person has failed to take advantage of it, someone else should be given a chance.	+4
30	Social class should make no difference whatsoever for prioritising health care. If people need treatment, they need treatment. How well off they are shouldn't come into it.	+4
25 ^a	People with dependants should not be given priority over people without dependants. A human life is a human life, I think it should be irrelevant how many dependants they've got.	+4
14 ^b	The age of the patient is important; if you were treating children rather than older people then you would have a longer improved life.	-4
20	People with dependants should be prioritised over people without dependants because their treatments would benefit others as well as the patient themselves.	-4
46	People who have already benefited from a lot of health care should take second place to people who have not used the health service as much.	-4
24	Whether or not people are currently working should be taken into account when we prioritise health care.	-4
3	People who live a healthy lifestyle should be prioritised because they would respond better to treatment.	-5
16 ^b	You should prioritise the younger age group, because they are still able to have children.	-5

a Denotes a statement which distinguishes factor 1 from factors 2 and 3 ($p < 0.05$).
 b Marks a significance level of $p < 0.01$.
 Consensus statements are shaded.

to divide the broad issue into sub-issues that could be described in discrete statements. Factor analysis showed a good deal of consensus that these issues should *not* be part of health-care priority setting, and we selected age and severity as the attributes to bring forward into the quantitative analysis. Future research investigating the views of the public around complex issues should consider using Q methodology in addition to more typical qualitative methods.

Diagrammatic questions

The presentation of the discrete choice and matching questions, including a detailed introductory explanation, was developed iteratively in focus groups. By far the most successful method was presentation of concepts of health (QoL), age, and health gains using diagrams. These diagrams were first explained by building them in small steps for respondents using an animated POWERPOINT

presentation (reproduced in Appendix 4), which was ultimately incorporated into a CAPI. The diagrams were then presented either as choice questions or as matching questions, examples of which are shown in *Figures 2* and *3*. Choice questions (e.g. *Figure 2*) present respondents with a one-off choice between option A and option B, which differ in terms of health gains, ages of patients and levels of QoL. The two options are presented both diagrammatically and descriptively in the accompanying text. Choice questions are explained in more detail in Chapter 4.

Matching questions (e.g. *Figure 3*) present respondents with a series of iterative choices where the numbers of people in Groups A and B are varied until a point of equivalence is reached. The same size of health gain is presented in Group A as in Group B for each set of iterative choices. These questions are explained in more detail in Chapter 5.

TABLE 3 Salient statements for factor 2

Number	Statement	Factor score
29 ^a	The quality of life of patients and their life expectancy are the most important things. The characteristics of patients like whether they are employed, or whether they have dependants, or what gender they are shouldn't matter.	+5
40	Priority should be given to preventive health care rather than always focusing on cure once people are ill.	+5
31	The amount of health and quality of life improvement is the most important. It's about getting the greatest benefit for the most people.	+4
30	Social class should make no difference whatsoever for prioritising health care. If people need treatment, they need treatment. How well off they are shouldn't come into it.	+4
23	Priority should be given to preventive health care especially education in schools about diet and lifestyle choices.	+4
44 ^a	It's no good saving lives if the quality of those lives is really bad. Some treatments are keeping people alive for too long. You've got to have a decent quality of life otherwise what's the point of being alive.	+4
18 ^a	There should be 'positive discrimination' towards people who are disadvantaged and in ill health because they've got a lot to contend with already.	-4
21 ^a	Older people deserve to be given priority. They have been paying into the NHS all their lives, they deserve to be able to draw on those resources when they need it.	-4
24	Whether or not people are currently working should be taken into account when we prioritise health care.	-4
9	People who smoke and drink pay enough in extra taxes to pay for their own health care.	-4
26 ^a	Poorer people should be given priority because they don't have the same opportunities to take care of their own health.	-5
28 ^a	Whether or not patients can contribute financially towards the cost of the treatment should be taken into account because it would allow you to treat more people who can't afford to 'go private'.	-5

a Marks a significance level of $p < 0.01$.
Consensus statements are shaded.

The questionnaire instrument was then piloted using cognitive interviews with 42 respondents (27 in Norwich and 15 in Newcastle) and was well received. Respondents followed the introduction and understood the questions. Their comments resulted in only minor adjustments to the wording and number of questions in each version of the questionnaire.

Survey sample and administration

The discrete choice and matching questions were part of a longer questionnaire (incorporating attitudinal questions and sociodemographics), which was administered face to face, using a CAPI, by interviewers from NatCen. The interview began with some basic demographic and household questions, following which an animated POWERPOINT presentation explained the meaning of the diagrams step by step. Next the respondents

answered six matching questions followed by eight discrete choice questions. There were four attitudinal questions, before some more detailed sociodemographic and health questions. On average, interviews lasted 41 minutes.

The survey was administered by NatCen to a random sample ($n = 587$) of the population in England from February to April 2007. The sample was generated by NatCen from the population of adults (aged 18 and over) living in England. Thirty addresses were selected from each of 40 postcode areas, which were stratified by Government Office Region (nine regions) and the proportion of manual/non-manual households. Within each household, only one adult was eligible for inclusion in the study. In households with more than one eligible adult present, interviewers randomly selected one interviewee. A total of 243 (41%) were male, the mean age of the whole sample being 52 years and, thus, females and older people are slightly over-represented.

TABLE 4 Salient statements for factor 3

Number	Statement	Factor score
40	Priority should be given to preventive health care rather than always focusing on cure once people are ill.	+5
23	Priority should be given to preventive health care especially education in schools about diet and lifestyle choices.	+5
13 ^a	Age shouldn't come into it, unless you're talking about children. Children's health should be given priority over adults.	+4
41	Health care should be based on <i>need</i> , not on social circumstances, or addiction or weight or smoking or anything else.	+4
30	Social class should make no difference whatsoever for prioritising health care. If people need treatment, they need treatment. How well off they are shouldn't come into it.	+4
31	The amount of health and quality of life improvement is the most important. It's about getting the greatest benefit for the most people.	+4
24	Whether or not people are currently working should be taken into account when we prioritise health care.	-4
20	People with dependants should be prioritised over people without dependants because their treatments would benefit others as well as the patient themselves.	-4
3	People who live a healthy lifestyle should be prioritised because they would respond better to treatment.	-4
17	For relatively minor conditions patients who are of working age should take priority over people who are retired.	-4
9 ^a	People who smoke and drink pay enough in extra taxes to pay for their own health care.	-5
46	People who have already benefited from a lot of health care should take second place to people who have not used the health service as much.	-5

a Marks a significance level of $p < 0.01$.
Consensus statements are shaded.

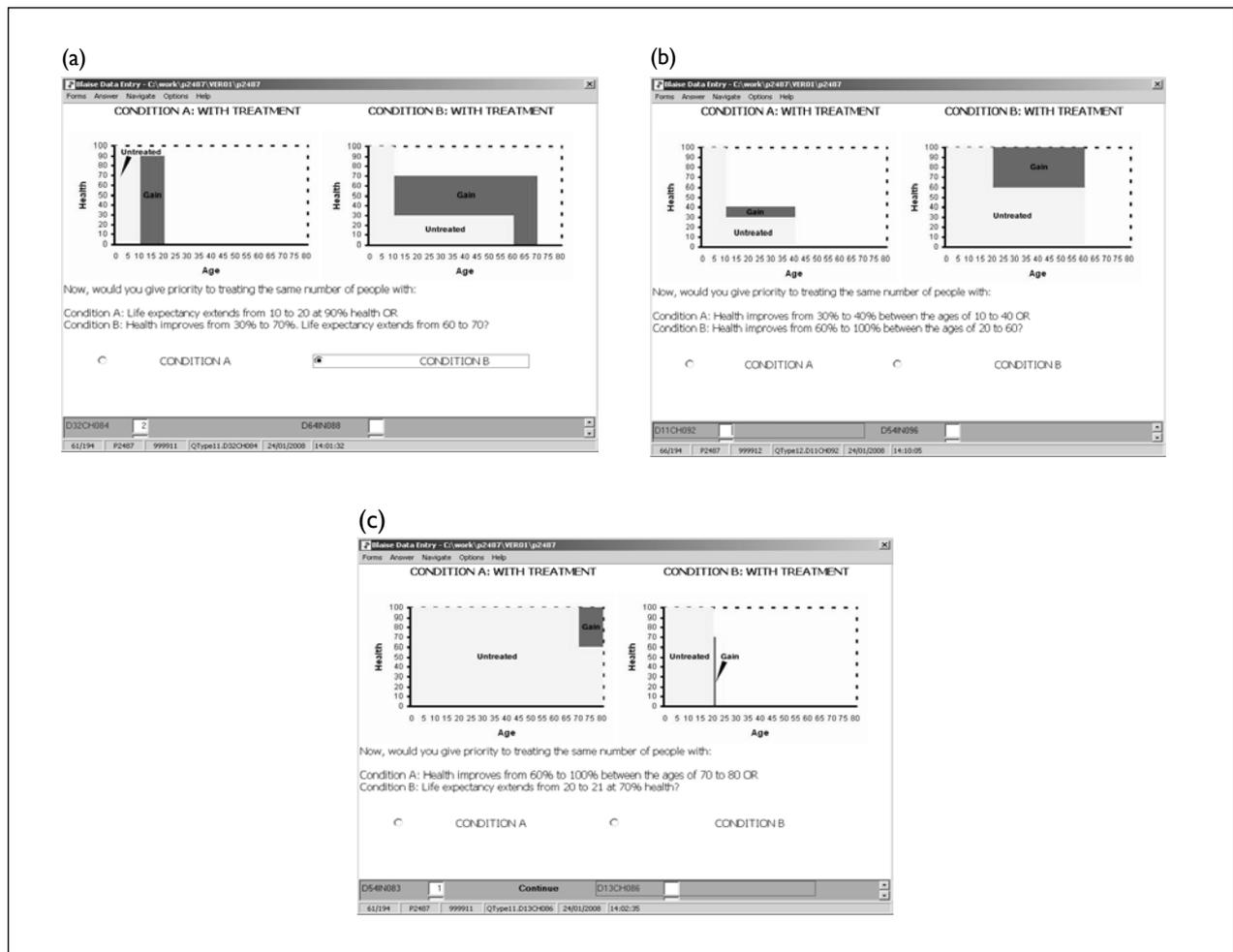


FIGURE 2 Examples of choice questions. (a) Screen 1: treatment for condition A results in a gain in life-years; treatment for condition B results in a gain in both quality of life (QoL) and length of life. (b) Screen 2: treatment for either condition A or condition B results in a gain in quality of life. (c) Screen 3: treatment for condition A results in a gain in quality of life; treatment for condition B results in a gain in length of life.

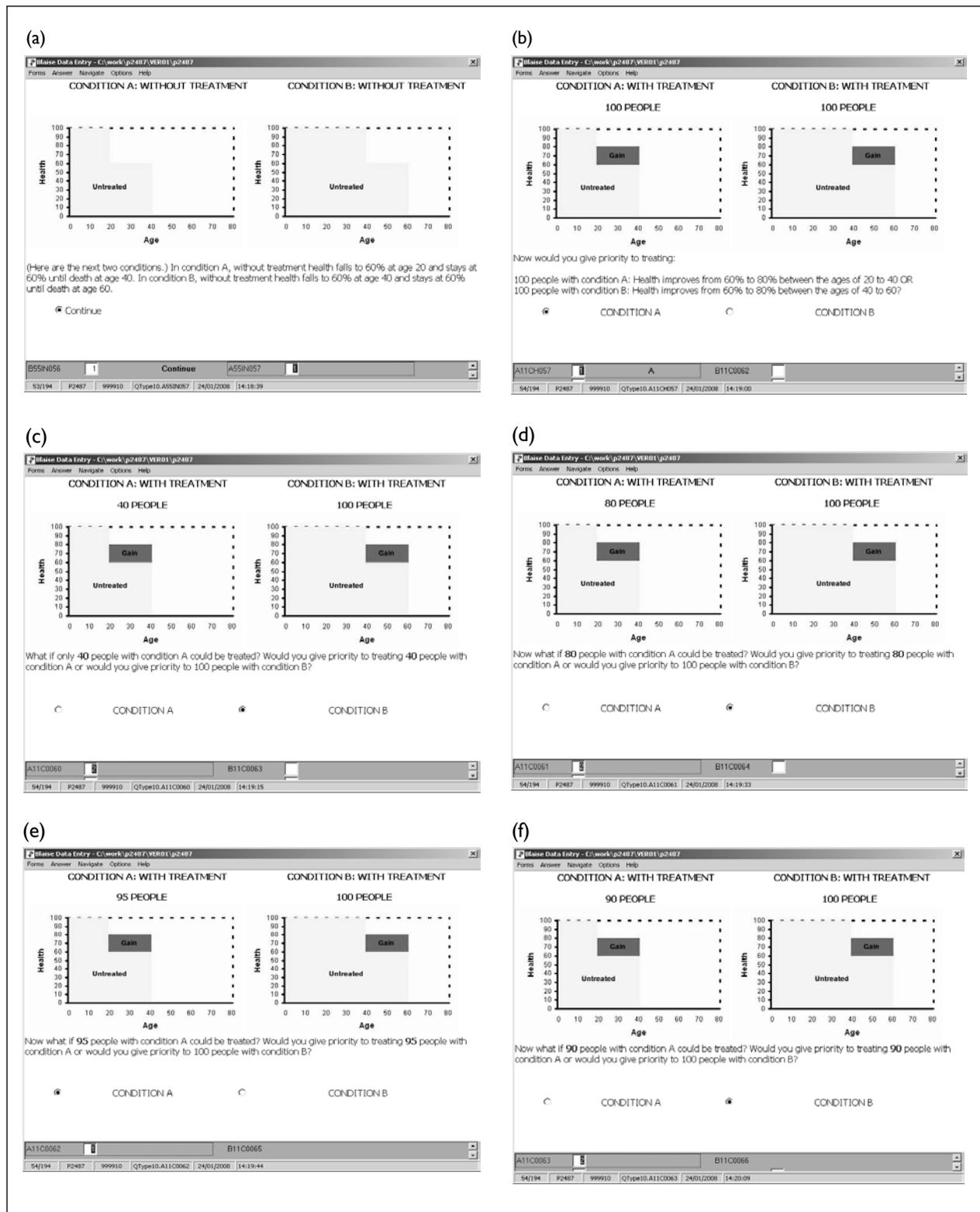


FIGURE 3 Example of iterations within a single matching question based on responses shown. (a) Screen 1: untreated scenarios. (b) Screen 2: 'with treatment' scenarios showing health gains – condition A is chosen by the respondent. (c) Screen 3: when 40 people with condition A or 100 people with condition B can be treated, the respondent chooses B. (d) Screen 4: when 80 people with condition A or 100 people with condition B can be treated, the respondent chooses B. (e) Screen 5: when 95 people with condition A or 100 people with condition B can be treated, the respondent chooses A. (f) Screen 6: finally, when 90 people with condition A or 100 people with condition B can be treated, the respondent chooses B.

Chapter 4

Discrete choice study

Basic design

In a discrete choice study, respondents are presented with a series of choice sets (usually pairwise). Each scenario in the set is defined according to some predefined attributes (ours being predetermined from the research reported in Chapter 3) and levels of such attributes. The attribute levels vary across scenarios and choice sets. In each choice set, the respondent is asked which scenario they prefer or would choose. Faced with a series of such choices, respondents essentially reveal how much weight they attach to each of the attributes, the actual weights being derived through statistical analysis of the data (see below).

Table 5 lists the attributes introduced in the preceding chapter along with the levels for each that were used in the discrete choice study.

The levels for the age variable were chosen to represent a range of stages of life: a very young child, a fully grown child and a young adult, followed by two or three further stages of adulthood into old age. The QoL if untreated variable again ranges from death through a series of better (though some still quite serious) states to 90%. The variables representing gains in life expectancy and QoL follow in large part from what was chosen for age and QoL if untreated. For example, to give a full number of life-years to a 1-year-old who would otherwise die, the gain in life expectancy would be 79, and to return someone in a health state valued at 60% back to full health, a gain of 40% would be offered. Some levels of the

life expectancy and QoL gain variables also reflect the desire to have some scenarios where people might get an improvement but not to full health or maximum life-years (the maximum here being 80).

Using these variables and imposing some assumptions (such as people being in full health before the onset of illness), we constructed diagrams of the sort shown in *Figure 2*. QALY gains (shown as a dark shaded area in the diagrams) are calculated from the five attributes listed in *Table 5*.

A full factorial design using the attributes listed in *Table 5* would have resulted in $(6)(7)(8)(4)(6) = 8064$ possible profiles but there are many combinations of levels on these attributes that result in implausible scenarios. For example, 'age at death' must be greater than 'age at onset' of illness, and the gain in life expectancy added to the age at death if untreated cannot exceed some reasonable maximum age, which was constrained to be 80 years. The full list of constraints is listed in Appendix 5. After imposing these constraints, 6572 of the possible 8064 profiles were implausible, leaving 1492 profiles (19% of the total).

The experimental design software SAS, which allows for constraints, was initially used to select choice sets from the 1492 profiles. This resulted in a design with over 200 profiles describing 'age at onset' as 1-year-olds and only 12 profiles describing 70-year-olds. At this stage the design was altered manually to improve the balance of questions about different age groups. Clearly, such a severely-reduced set of available profiles, together with the manual alteration of the design to achieve greater

TABLE 5 Attributes and levels

Description of attribute	Levels
Age at onset (years)	1, 10, 20, 40, 60, 70
Age at death if untreated (years)	1, 10, 20, 40, 60, 70, 80
Gain in life expectancy (years)	0, 1, 5, 10, 20, 40, 60, 79
QoL if untreated (represented as %)	0, 30, 60, 90
Gain in QoL with treatment (%)	0, 10, 20, 40, 70, 100
QoL, quality of life.	

balance, negatively affect the design properties that are desirable in discrete choice experiments (DCEs). Nevertheless, before administering the questionnaire, data were simulated to ensure that a model could be estimated on the basis of the amended design.

Functional form and empirical approach

Response data in DCEs are modelled within a random utility framework of the general form:

$$U = V + \varepsilon \quad (1)$$

in which utility, U , is separated into parts which are explainable, in this case V , and unexplainable, ε . In this study, we are concerned with estimating V , which represents an underlying continuous and latent variable which is nevertheless unobservable.

If the standard QALY model is true, V would simply be a function of QALYs. If individuals are concerned about other characteristics, then these will also be part of the utility function. In this simple case we assume that utility is a function of age at onset (AO), age at death without treatment (AD), QoL lost without treatment (QL) and QALYs gained from treatment ($QALY$). This gives:

$$V = f(AO, AD, QL, QALY) \quad (2)$$

Quality of life lost (QL) is transformed from a variable in *Table 5*, by subtracting 'quality of life if untreated' from 1. This is done in order to facilitate the log transformation required below – with some scenarios involving instant death, and thus a QoL of zero, which could not otherwise have been log transformed. Note, therefore, that when interpreting this variable, the larger the 'quality of life lost' at the onset of illness, the more 'severe' the health state.

The first two terms on the right-hand side of (2) detect age effects, QL detects severity and $QALY$ is the health gain. If expressed as an additive function, this would mean that gains in utility could be experienced even if QALY gains were zero, i.e. age and severity would have effects on utility irrespective of whether or not QALY gains are incurred. The alternative, therefore, was to use a multiplicative form of the utility function. The QALY itself is a multiplicative function of life-years and QoL gained. By extension, in a multiplicative form of the above function, utility is derived from

QALYs multiplied by the magnitudes of the other variables, ensuring that with zero QALYs gained there is a zero impact on utility. Empirically, the multiplicative models presented below consistently outperformed those based on an additive functional form, which were also investigated. Although, it may appear that a multiplicative model of the form $QALYs \times AGE \times SEVERITY$, with just one age-related variable, would make more theoretical sense, we took the more pragmatic view that this would leave too much riding on the 'age at onset' variable in terms of what respondents might be thinking about in relation to age, and so we included age at death as well. In addition, this was the best performing model empirically, which, it could be argued, is important for estimating weights.

If we assume such a multiplicative underlying model, we may use a log-linear model of the form:

$$\log(V) = \beta_1 \log(AO) + \beta_2 \log(AD) + \beta_3 \log(QL) + \beta_4 \log(QALY) \quad (3)$$

This is a standard log-linear utility function where the β s are parameters to be estimated. Given that discrete choice response data are based on choices over alternative combinations of the dependent variables, then, assuming that the β parameters are identical across all individuals, a simple model of the following form can be estimated:

$$\Delta \log(V) = \beta_1 (\Delta \log(AO)) + \beta_2 (\Delta \log(AD)) + \beta_3 (\Delta \log(QL)) + \beta_4 (\Delta \log(QALY)) \quad (4)$$

where Δ represents the differences in levels of any given attribute reflected in the pairwise choices presented.

Equation (4) was estimated using a logit model (see 'simple' model under Discrete choice results), allowing for clustering of the individual standard errors to account for the fact that each individual responded to several questions. Essentially, this amounts to a conditional logit model.

A flexible functional form (referred to below as the 'powered' model) was also specified to allow for any non-linear relationships between choice and the included variables. While the standard log model allows for non-linearities, these functions are monotonic, this restriction possibly being too strong and resulting in biased estimates. Including higher order terms, in a fashion analogous to the popular translog model,⁵⁶ allows for modelling of more complex non-linearities. The use of such

flexible functional forms provides more robust estimation of the coefficients and reduces the potentially confounding problems of omitted variable bias. This also allows us to investigate the non-linearities detected in the matching data (see Chapter 5). All models were tested using the Akaike and Bayesian information criteria in order to aid model selection. Those which performed best on these criteria are reported below.

Discrete choice results

A total of 587 respondents yielded 4696 useable responses to the discrete choice questions. The estimated models are shown in *Table 6*. We did also take a more conventional econometric approach of converting covariates into categorical variables (i.e. sets of dummy variables) within the basic multiplicative framework. However, these models did not perform as well as those reported and are more challenging for calculation of weights.

The coefficients for the simple model suggest that increasing age at onset reduces the probability of choice, as does increasing the age at death, although the former is not statistically significant. This suggests that the young are preferred to the old. The coefficient on QoL lost is also negative, suggesting that as the health state is more severe the respondents are less likely to choose that group. Finally, the impact of QALY is positive, as we would

expect. The impact of severity, as it appears in the regression results, appears to contradict earlier literature. However, a more accurate picture of the impact of all of the variables is provided through examination of the shapes of the functional forms, which, for the simple model, are given in *Figures 4–7*. These diagrams show the shapes of the functional relationships when measured on a single scale, as represented by the ‘predicted utility’ axes. Here, it can be seen that the general impact of the QoL lost variable on utility is very small, the inference being that the severity–utility relationship is essentially flat and nearly so for age.

The equivalent diagrams for the powered models are shown in *Figures 8–11*. Once again, the sign of the coefficients in the regression model inform us of the direction of impact on utility for age at onset (individuals who are aged around 10–40 years are slightly preferred to the very young and the very old), age at death (slight preferences to save those who die young and those who will die old rather than the middle aged) and severity (a preference for individuals with lower severity, with the maximum at 0.4, after which predicted utility slopes downwards showing less preference for purely life-saving interventions). However, *Figures 8–10* demonstrate, once again, the relationship between each of these variables and utility to be essentially flat. This is reinforced by the lack of statistical significance on some of the coefficients, especially for age at onset and, this time, QoL

TABLE 6 Simple and powered models

	Simple model			Powered model		
	Coefficients	Standard error	p-value	Coefficients	Standard error	p-value
log AO	−0.02	0.022	0.304	−0.31	0.264	0.240
(log AO) ²				0.24	0.164	0.151
(log AO) ³				−0.04	0.025	0.107
log AD	−0.07	0.034	0.034	1.28	0.314	0.000
(log AD) ²				−0.76	0.175	0.000
(log AD) ³				0.11	0.025	0.000
log QL	−0.14	0.037	0.000	−0.64	0.372	0.085
(log QL) ²				−0.43	0.489	0.381
(log QL) ³				−0.09	0.149	0.559
log QALY	0.75	0.033	0.000	0.45	0.054	0.000
(log QALY) ²				−0.03	0.028	0.237
(log QALY) ³				0.03	0.007	0.000

AD, age at death without treatment; AO, age at onset; QALY, QALYs gained from treatment; QL, quality of life lost without treatment.

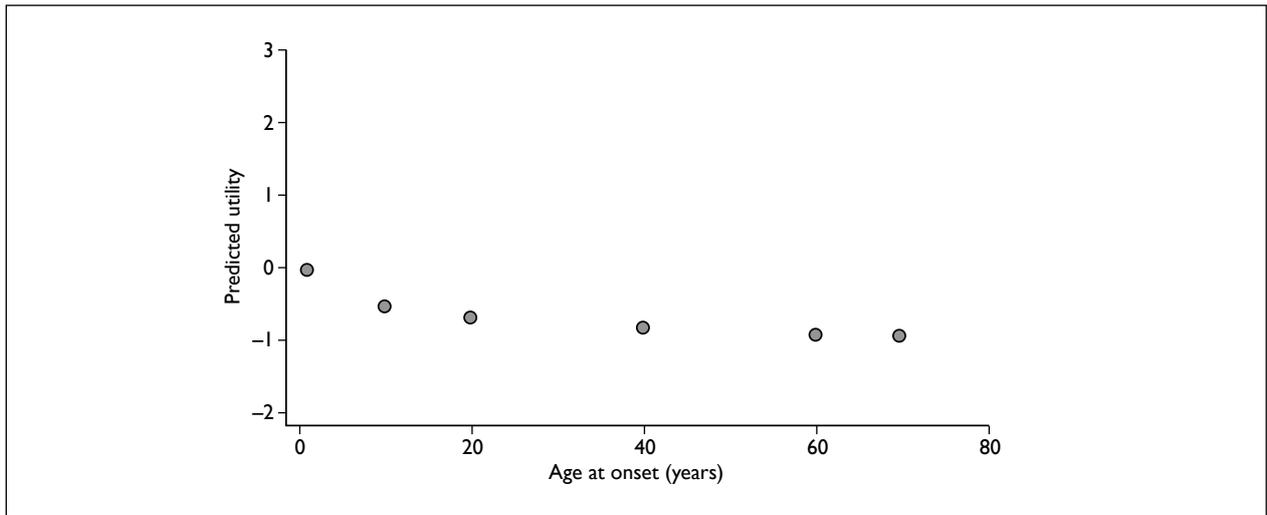


FIGURE 4 Age at onset vs utility.

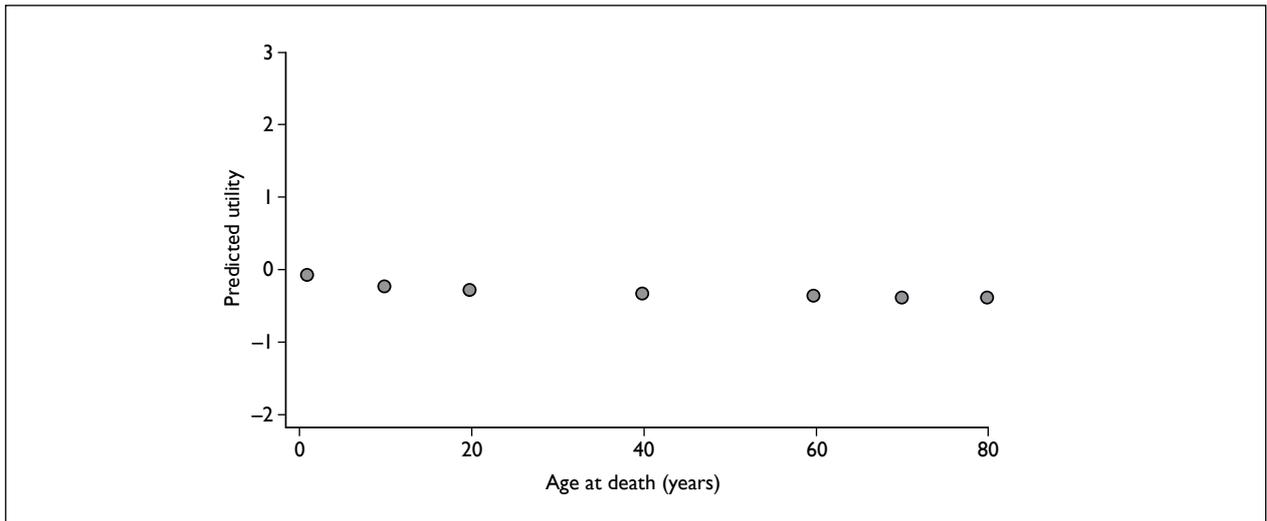


FIGURE 5 Age at death vs utility.

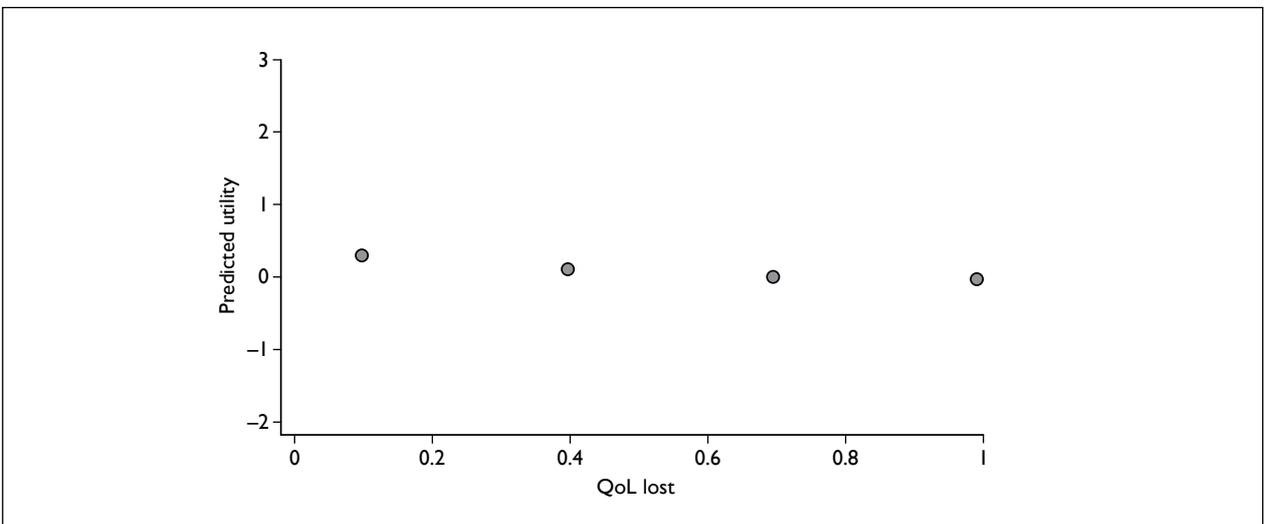


FIGURE 6 Quality of life (QoL) lost vs utility.

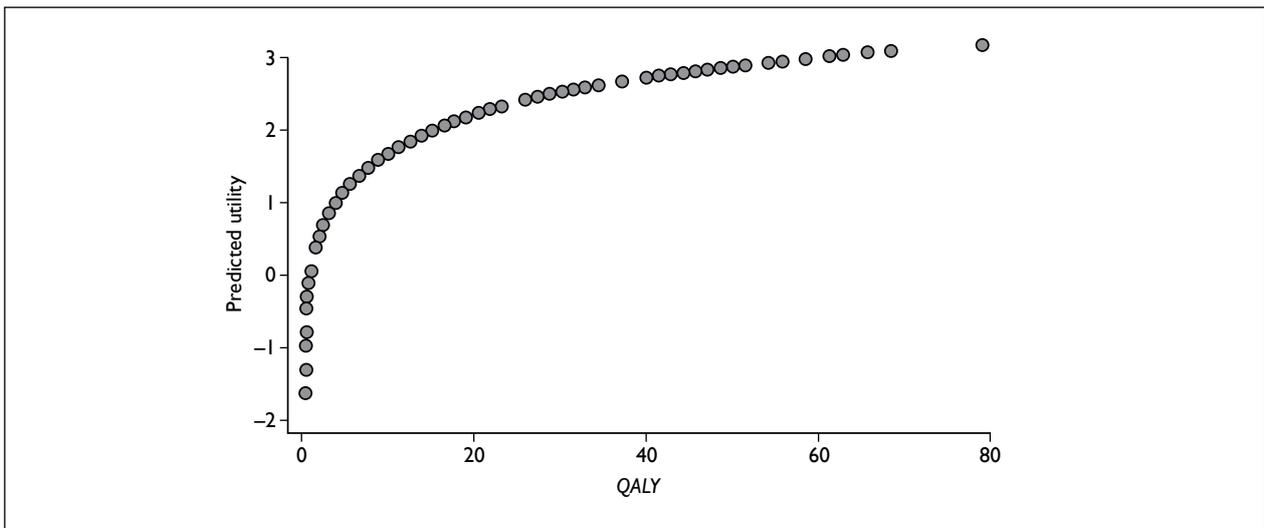


FIGURE 7 QALY vs utility.

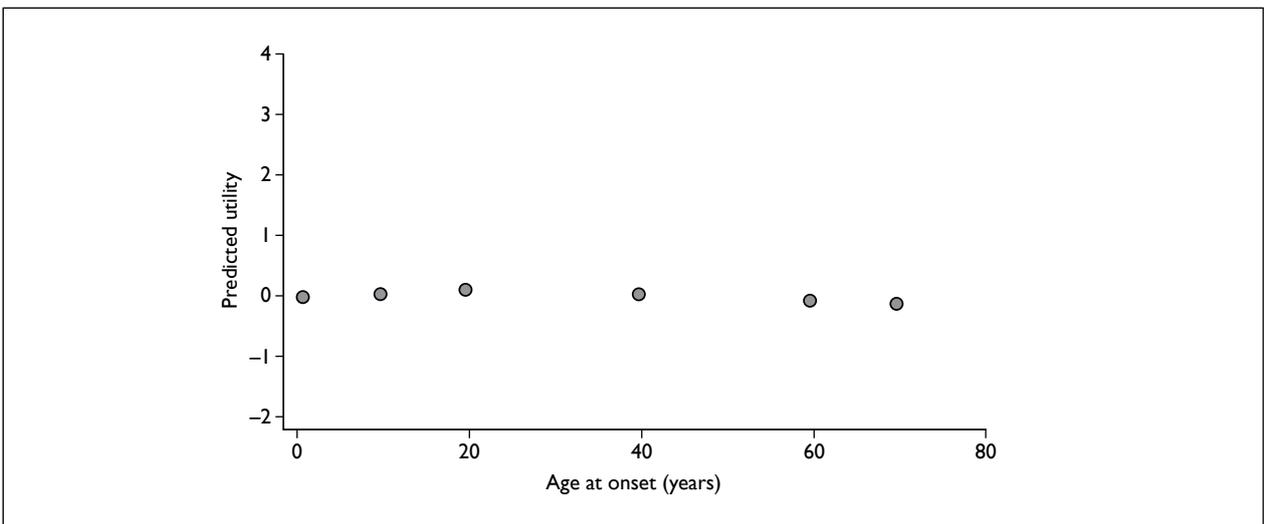


FIGURE 8 Age at onset vs utility (powered functions).

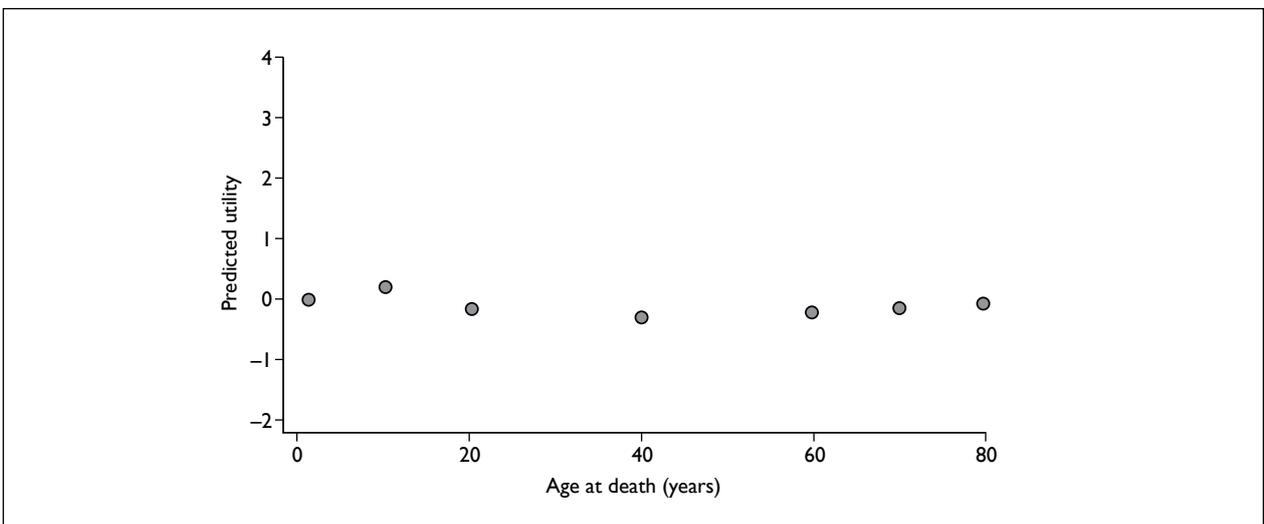


FIGURE 9 Age at death vs utility (powered functions).

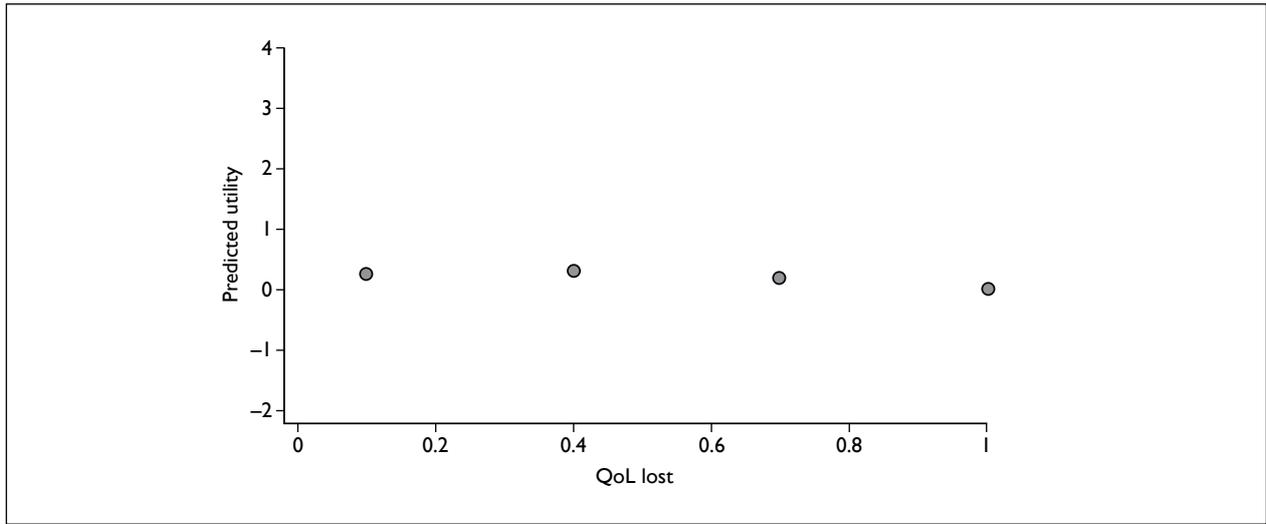


FIGURE 10 Quality of life (QoL) lost vs utility (powered models).

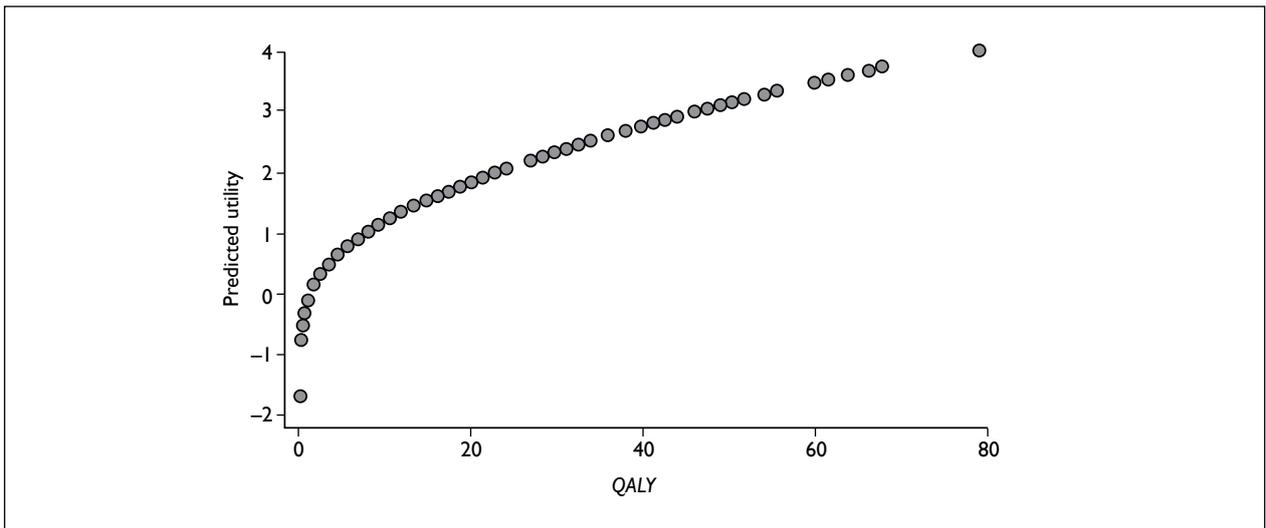


FIGURE 11 QALY vs utility (powered model).

lost. Nevertheless, weights were calculated using the point estimates from the powered regression, regardless of statistical significance. This approach is used for a number of reasons: first, the model is the best performing in diagnostic tests; second, we cannot rule out joint significance; and third, and most importantly, when calculating weights from any model, the point estimates from that model (regardless of which approach is used) are our most informed values. While it is possible to use hypothesis testing to determine whether our estimates are significantly different to zero, this does not provide better information than the point estimates. Also, because of what was revealed by the directions of the coefficients in the regression models, the results on severity are further

investigated (see Further investigation of severity). *Figure 11* shows that increasing QALYs are always preferred. Concavity is still observed, but less so than in the simple model.

Weighting QALYs using discrete choice data

Two novel ways of estimating weights from the above models have been devised, one based on predicted probabilities of choice and the other based on a compensating variation approach. As indicated above, for each, the weights presented are based on the powered model only.

Predicted probability of choice approach

The weights are estimated using the estimated probability of choice, with a base case of $AO = 40$, $AD = 60$, $QL = 0.7$, and $QALY = 4$ (i.e. individuals fall ill at 40, lose 0.7 of their QoL, will die at 60 without treatment, and are then given four QALYs with treatment). The way in which the four QALYs have been allocated is unspecified in this model. This choice of a base case is challenging for interpretation of subsequent weights, in that it may be thought better to choose an extreme position and then measure weights for every other scenario in one direction relative to that. However, using an extreme as a base case is problematic too; the most obvious example being use of the highest age group from which it is not possible to gain any QALYs. The choice of a four-QALY gain arose to reflect a reasonable-sized gain and also to correspond with the four QALY gains which were offered to respondents in the matching study (see Chapter 5).

In order to calculate the weights we compare our base case to an alternative scenario. We then vary the number of QALYs being offered in the alternative scenario until the estimated probability of choosing the base case equals 0.5 (i.e. the individual is indifferent between the two scenarios). This is more easily demonstrated by using an example. We have our base case, $AO = 40$, $AD = 60$, $QL = 0.7$ and $QALY = 4$. We now take a scenario for which we want to find a weight, for example $AO = 1$, $AD = 1$ and $QL = 1$ (i.e. individuals fall ill and die at age 1), this being comparison 1 in Table 7. We adjust the number of QALYs from treatment available to the comparison group until the probability of choosing the base case equals 0.5. This probability to calculated as:

$$Pr_{BASECASE} = \frac{\exp(\log(u_{BASECASE}))}{\exp(\log(u_{BASECASE})) + \exp(\log(u_{COMPARISON}))}$$

where $\exp(\log(u_i))$ is the predicted utility for the i th choice from the powered model regression results in Table 6. In this case the probability of choosing the base case is 0.5 when the comparison group is offered 4.1 QALYs. The weight itself is found by taking the ratio of the QALYs offered in both cases, so:

$$weight = \frac{QALY_{COMPARISON}}{QALY_{BASECASE}}$$

which in this case gives $4.1/4 = 1.025$. This demonstrates that the base case is slightly preferred to the alternative, with one QALY to the base case being equal to 1.025 QALYs to the comparison. This process is repeated for each comparison scenario to generate weights. More generally, we try to illustrate this in Figure 12. Weights closer to 0 show a stronger preference for the comparison, weights equal to 1 show indifference between the two groups and weights greater than 1 show preference for the base case.

The weights from this procedure are given in Table 7. (In this table, the variables given earlier as percentages are now presented on a 0–1 scale in line with the more common representation in the QALY literature. Percentages were used earlier because that is how QoL was presented to survey respondents.) While these represent weights arising from assessing the respective comparator group against the base case, it is possible to generate weights for comparing different comparison groups. For example, if we wished to compare scenarios 5 and 8, we could do this indirectly by comparing the weights given in Table 7. The weight for comparison 5 is 0.79 and the weight for comparison 8 is 0.725 – from this we can conclude that scenario 8 is preferred to scenario 5. Alternatively, we could take the ratio of these weights, to generate a new weight for scenario 8 of 0.92 ($0.725/0.79$), which demonstrates that 0.92

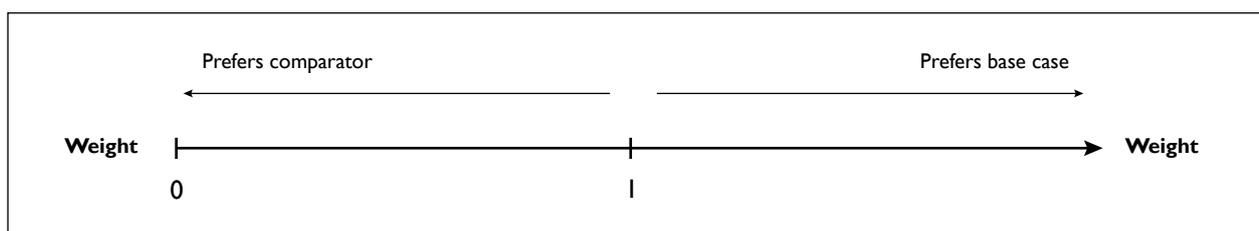


FIGURE 12 Preferences and weights in the discrete choice study.

of a QALY to scenario 8 is worth one QALY to scenario 5.

The weights presented in *Table 7* are sometimes difficult to interpret because a number of factors vary at once when comparing scenarios. We can see that most weights are less than 1, suggesting that the comparison scenarios used here are generally preferred to the base case. The only exceptions are comparisons 1, 9, 11, 14 and 15 (and comparison 12 which is the same as the base case). Apart from comparison 1, these scenarios are towards the top end of age at death and are quite severe.

For those scenarios which are purely life saving, where age at onset equals age at death (comparisons 1, 6, 11 and 14), we see that there is a preference for those with an age at onset and age at death of 10 (comparison 6). The weight for this scenario is 0.62. The weights for the older groups (comparisons 11 and 14) are greater than 1, indicating a preference for the base case. The weight for the youngest group (comparison 1) is also greater than 1, suggesting that individuals do not choose to treat those who are very young. These results suggest a non-linear relationship

between age and weights, with the most preferred being the younger, but not youngest, groups.

For severity we can see that, for the same age at onset and age at death, as QoL lost decreases, moving from 0.7 to 0.1, the weight moves towards 0. The less severe comparisons (where QoL lost equals 0.1) have a weight of less than 1, showing that they are the preferred groups. This can be seen for comparisons 2 and 3, 4 and 5, 7 and 8, 9 and 10, 12 and 13, and 15 and 16.

It is difficult to make generalisations about age at onset and age at death using *Table 7*, where both variables often vary at the same time. An alternative approach to presenting the weights is to hold one of the factors constant and allow the others to vary. This is the approach used in *Tables 8* and *9*.

In *Table 8*, QoL lost is fixed at 0.7. This clearly shows that as age at onset increases the weights move towards 0 up to age 20 but then move towards 1. Weights for age at death move towards 1 up to age 40, but then fall again. This shows that there is a clear preference for treating people

TABLE 7 *Weights based on powered model*

Base	Age at onset 40	Age at death 60	QoL lost 0.7	QALYs gained 4	Weight
Comparison					
1	1	1	1	4.1	1.025
2	1	10	0.7	1.85	0.463
3	1	10	0.1	1.52	0.38
4	1	20	0.7	3.8	0.95
5	1	20	0.1	3.16	0.79
6	10	10	1	2.48	0.62
7	10	20	0.7	3.5	0.875
8	10	20	0.1	2.9	0.725
9	10	40	0.7	4.6	1.15
10	10	40	0.1	3.95	0.9875
11	40	40	1	6.15	1.54
12	40	60	0.7	4	1
13	40	60	0.1	3.38	0.845
14	70	70	1	6.45	1.61
15	70	80	0.7	4.2	1.05
16	70	80	0.1	3.55	0.89

QoL, quality of life.

TABLE 8 Weights by age at onset and age at death

Age at death	Age at onset					
	1	10	20	40	60	70
1						
10	0.47					
20	0.95	0.87				
40	1.25	1.16	1.04			
60	1.09	1.00	0.90	1.00		
70	0.95	0.88	0.78	0.86	1.09	
80	0.80	0.74	0.65	0.73	0.93	1.05

who fall ill at ages between 10 and 40 rather than the very old or the very young, and a preference for treating those who will die either very young or very old. In this, and similar tables to follow, blank cells are simply representative of unfeasible comparisons.

In *Table 9*, age at death is fixed at 60 and the other attributes are allowed to vary. Again we can see that the weights for age at onset are quadratic, starting close to 1 for an age at onset of 1, moving towards 0 up to age at onset of 20 before moving towards 1 as age at onset increases further. For severity we can see a preference for treating those who are closer to the middle of the severity range. The weights are nearest 0 for a severity of 0.4 and closer to 1 for severity scores of 0.1 (least severe) and 1 (most severe).

Compensating variation approach

Another approach to calculating the relative weights attached to different types of QALYs, or beneficiaries of QALYs, is to use the Hicksian compensating variation approach to welfare measurement. The method for calculating the compensating variation using discrete data is due to Small and Rosen⁵⁷ and was introduced to the

health economics literature in the context of DCEs by Lancsar and Savage.⁵⁸

In general, the compensating variation (CV) is calculated by valuing in monetary terms the change in expected utility due to a policy change (e.g. change in price or quality of a good/service, or, in our study, a change in health state) as the change in income required to return the individual to their initial level of utility, that is, to compensate them for the change.

The CV for discrete choice data takes the following form:

$$CV = -\frac{1}{\lambda} \left[\ln \sum_{j=1}^J e^{V_j^0} - \ln \sum_{j=1}^J e^{V_j^1} \right] \quad (5)$$

where V_j^0 and V_j^1 are the value of the indirect utility function for each choice option j before and after the policy change respectively; J is the number of options in the choice set; and λ is the marginal utility of income, or its proxy.

While a monetary value is the most convenient to turn the change in expected utility into a common metric, in fact any quantitative metric could be used as the numeraire. In the current study, instead

TABLE 9 Weights by age at onset and severity

Severity	Age at onset				
	1	10	20	40	60
0.1	0.925	0.850	0.750	0.840	
0.4	0.875	0.805	0.713	0.800	
0.7	1.088	1.000	0.900	1.000	
1					1.625

of using income it is possible to calculate the CV for a move from one health state to another using the marginal utility of QALY gains as the numeraire. That is, we calculate the number of additional QALYs required to be given or taken away in the new health state to equalise the utility derived from the base and new health states.

Like the probability-of-choice approach described previously, the CV approach can be used to value changes in, and create weights for, whole scenarios (i.e. entire health states). In addition, the CV can also be calculated to value and derive weights for individual characteristics such as age at onset, age at death and severity. To calculate both types of weights, we value changes from an initial health state, namely the same reference case as used in the calculation of the probability weights: $AO = 40$, $AD = 60$, $QL = 0.7$ and $QALY = 4$.

From this initial health state we can calculate the CV, measured in numbers of QALYs, associated with a move to a new health state, described by new levels on the above attributes. This can involve an entire scenario or the valuation of each

attribute one at a time. We use these CV measures to calculate weights for entire health states and for individual characteristics that describe the health states, by taking the ratio of the total QALYs required in the new health state to equalise utility and the original number of QALYs in the base model. This welfare theoretic approach also allows investigation of the strength of preference, as indicated by the magnitude of the CV.

As with the probability-of-choice approach, the CV and relative weights reported in this section used the conditional logit results for the powered model reported in *Table 6*. The CV and relative weights for the entire health states described in *Table 7* are reported in *Table 10*. In each of these, QALY gains are held constant at 4.

Before discussing the results in *Table 10*, it is important to note that 'QALY gain with treatment' represents the number of QALYs in the health-state scenarios presented to respondents in the survey. This is not the same as the 'QALY gain' in the probability-of-choice tables, which is the result of the calculation of the number of QALYs required to

TABLE 10 Weights based on compensating variation approach (version 1)

Base	Age at onset 40	Age at death 60	QoL lost 0.7	QALY gain with treatment 4	CV	Weight per scenario
Comparison						
1	1	1	1	4	0.0224	1.0056
2	1	10	0.7	4	-0.5788	0.8553
3	1	10	0.1	4	-0.7280	0.8180
4	1	20	0.7	4	-0.0435	0.9891
5	1	20	0.1	4	-0.1787	0.9553
6	10	10	1	4	-0.3650	0.9087
7	10	20	0.7	4	-0.1097	0.9726
8	10	20	0.1	4	-0.2468	0.9383
9	10	40	0.7	4	0.1190	1.0297
10	10	40	0.1	4	-0.0112	0.9972
11	40	40	1	4	0.3591	1.0898
12	40	60	0.7	4	0	1
13	40	60	0.1	4	-0.1339	0.9665
14	70	70	1	4	0.4050	1.1013
15	70	80	0.7	4	0.0390	1.0098
16	70	80	0.1	4	-0.0937	0.9766

CV, compensating variation; QoL, quality of life.

equalise the probability of choosing the base case and alternative health states.

Both the CV and weights are interpreted relative to the reference case. The same reference case is included as comparison 12, for which the CV is 0, as would be expected since the health state has not changed, and the weight is 1. In *Figure 13*, relative to this base case, a negative CV indicates that the new state is preferred to the base case and that individuals are 'willing to pay' to secure the new health state – see for example, comparison 2. The consequent weight of 0.855 indicates that a QALY given in the base state is worth only 0.855 of a QALY given in the new state. A positive CV indicates that individuals require compensation (in terms of being given additional QALYs) in the new health state – see for example, comparison 1. A weight of greater than 1 indicates preference for the base state.

For the four comparison health states that involve instant death without treatment (comparisons 1, 6, 11 and 14) there is a preference for giving the QALYs to the base case, with the exception of comparison 6 which favours the alternative. This is in accordance with the underlying utility values derived from each health state. That is, the utility derived from the base case is greater than the utility derived from the alternatives with the exception of comparison 6 (a 10-year-old who dies instantly without treatment), which yields a higher utility than the base case.

Still looking at the magnitude of the CV relative to the base case, respondents were willing to pay the largest amount (in terms of QALYs) for comparison 3, namely a health state in which the individual becomes unwell at age 1, without treatment QoL drops by 90% and they die at age 10, whereas with treatment they gain four QALYs. The least preferred beneficiary of four QALYs is someone

who becomes unwell at age 70 and without treatment dies instantly (comparison 14).

Welfare measures and relative weights can also be calculated for the same complete health states described in *Table 10*, but where QALYs, in addition to the other attributes, are allowed to change between the initial and new health states. These results are presented in *Table 11*.

Again, a weight greater (less) than 1 indicates a preference for the base (alternative) health state. Looking at these results in aggregate, the number of QALYs gained is driving the preference between the base and comparison health states. Without exception, for QALY gains in the new health state of 10, 15 and 30, the new health state is always preferred to the base (which only contains four QALYs). For all comparison health states involving only one QALY gain, the base case (containing four QALYs) is always preferred. For comparison health states with two QALY gains, the base case is generally preferred to the comparison case, except when those two QALYs are given to 1-year-olds whose QoL drops by 0.1 and who die at age 10, or 1-year-olds whose QoL drops by 0.7 and who die at age 20.

The CV and relative weights associated with the individual attributes (age at onset, age at death and severity) are reported in *Table 12*. These calculations used the same initial base health state of age at onset of 40, age at death of 60, QoL lost of 0.7 and four QALYs. The CV and weights per attribute level were calculated by changing one attribute at a time in the new health state, holding all else constant.

Age at onset of 20 and 40 appear to be the most important, which could relate to considerations of the ages at which individuals are most productive to society (and also more likely to have

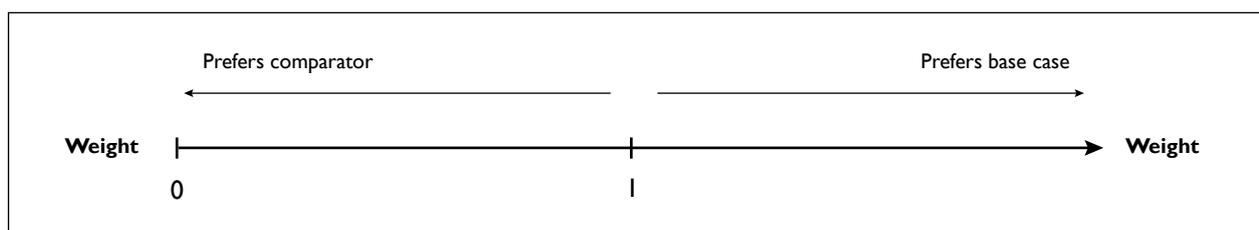


FIGURE 13 Preferences and weights in the discrete choice study.

TABLE II Weights based on compensating variation approach (version II)

Base	Age at onset 40	Age at death 60	QoL lost 0.7	QALY gain with treatment 4	CV	Weight per scenario
Comparison						
1	1	1	1	1	0.8568	1.2142
2	1	10	0.7	1	0.3853	1.0963
3	1	10	0.1	1	0.2642	1.0660
4	1	20	0.7	1	0.8065	1.2016
5	1	20	0.1	1	0.7022	1.1756
6	10	10	1	1	0.5561	1.1390
7	10	20	0.7	1	0.7557	1.1889
8	10	20	0.1	1	0.6491	1.1623
9	10	40	0.7	1	0.9299	1.2325
10	10	40	0.1	1	0.8312	1.2078
11	40	40	1	1	1.1082	1.2771
12	40	60	0.7	1	0.8398	1.2100
13	40	60	0.1	1	0.7370	1.1842
14	70	70	1	1	1.1417	1.2854
15	70	80	0.7	1	0.8695	1.2174
16	70	80	0.1	1	0.7680	1.1920
17	1	1	1	2	0.4895	1.1224
18	1	10	0.7	2	-0.0460	0.9885
19	1	10	0.1	2	-0.1813	0.9547
20	1	20	0.7	2	0.4316	1.1079
21	1	20	0.1	2	0.3122	1.0780
22	10	10	1	2	0.1462	1.0366
23	10	20	0.7	2	0.3733	1.0933
24	10	20	0.1	2	0.2517	1.0629
25	10	40	0.7	2	0.5741	1.1435
26	10	40	0.1	2	0.4601	1.1150
27	40	40	1	2	0.7821	1.1955
28	40	60	0.7	2	0.4699	1.1175
29	40	60	0.1	2	0.3519	1.0880
30	70	70	1	2	0.8215	1.2054
31	70	80	0.7	2	0.5041	1.1260
32	70	80	0.1	2	0.3874	1.0969
33	1	1	1	10	-0.9594	0.7602
34	1	10	0.7	10	-1.6568	0.5858
35	1	10	0.1	10	-1.8256	0.5436
36	1	20	0.7	10	-1.0374	0.7407
37	1	20	0.1	10	-1.1961	0.7010
38	10	10	1	10	-1.4122	0.6469

TABLE II Weights based on compensating variation approach (version II) (continued)

Base	Age at onset 40	Age at death 60	QoL lost 0.7	QALY gain with treatment 4	CV	Weight per scenario
39	10	20	0.7	10	-1.1153	0.7212
40	10	20	0.1	10	-1.2754	0.6811
41	10	40	0.7	10	-0.8442	0.7890
42	10	40	0.1	10	-0.9992	0.7502
43	40	40	1	10	-0.5535	0.8616
44	40	60	0.7	10	-0.9859	0.7535
45	40	60	0.1	10	-1.1437	0.7141
46	70	70	1	10	-0.4971	0.8757
47	1	1	1	15	-1.6162	0.5960
48	1	10	0.7	15	-2.3566	0.4109
49	1	10	0.1	15	-2.5338	0.3666
50	1	20	0.7	15	-1.6997	0.5751
51	1	20	0.1	15	-1.8691	0.5327
52	10	10	1	15	-2.0985	0.4754
53	10	20	0.7	15	-1.7829	0.5543
54	10	20	0.1	15	-1.9535	0.5116
55	10	40	0.7	15	-1.4924	0.6269
56	10	40	0.1	15	-1.6588	0.5853
57	40	40	1	15	-1.1777	0.7056
58	40	60	0.7	15	-1.6446	0.5889
59	40	60	0.1	15	-1.8132	0.5467
60	1	1	1	30	-3.2013	0.1997
61	1	10	0.7	30	-4.0036	-0.0009
62	1	10	0.1	30	-4.1926	-0.0482
63	1	20	0.7	30	-3.2930	0.1768
64	1	20	0.1	30	-3.4780	0.1305
65	10	10	1	30	-3.7265	0.0684
66	10	20	0.7	30	-3.3840	0.1540
67	10	20	0.1	30	-3.5696	0.1076
68	10	40	0.7	30	-3.0648	0.2338
69	10	40	0.1	30	-3.2481	0.1880
70	40	40	1	30	-2.7139	0.3215
71	40	60	0.7	30	-3.2325	0.1919
72	40	60	0.1	30	-3.4171	0.1457

CV, compensating variation; QoL, quality of life.

TABLE 12 Compensating variations per attribute

Attribute	Level	CV	Weight
Age at onset	1	0.0688	1.0172
	10	0.0043	1.0011
	20	-0.0862	0.9784
	40	0	1
	60	0.1741	1.0435
	70	0.2675	1.0669
Age at death	1	-0.3173	0.9207
	10	-0.6568	0.8358
	20	-0.1141	0.9715
	40	0.1148	1.0287
	60	0	1
	70	-0.1108	0.9723
	80	-0.2438	0.9390
Severity (QoL loss)	1	0.2517	1.0629
	0.7	0	1
	0.4	-0.1730	0.9568
	0.1	-0.1339	0.9665

CV, compensating variation; QoL, quality of life.

dependants), followed by the very young, with least importance given to the eldest two age groups of 60- and 70-year-olds. Higher preference is given to those who die young (aged 10, 1 and 20).

When interpreting the severity weights, we first note that the base case of QoL lost = 0.7 is one of the middle severity levels, so movements from this base case can represent an improvement in QoL (a loss of only 0.4 or 0.1 rather than 0.7 QoL) for which respondents were willing to pay or a detriment (QoL lost = 1) for which respondents required compensation. As movements from the base are in both directions, absolute values are used. Relative to the base case, again, more severe states are less preferred, as would be expected given the results reported above. In particular, moving from the base to the worst level of severity had the largest impact in terms of the CV; that is, respondents would have to be compensated a relatively large amount in absolute value for a health state that involved instant death. Willingness to pay (in terms of QALYs) to improve the initial level of severity – that is, to lose only 10% or 40% of QoL rather than the base case of 70% was similar across these two levels, producing similar weights (0.9567 and 0.9665 respectively).

Further investigation of severity

In the above regression analyses, the severity variable (QoL lost) appeared to behave in the opposite direction to how it had been weighted in previous studies. Although the relationship is close to being flat, this different result is nevertheless reflected to some extent in the regressions and in the weights presented in the previous section. This issue was, therefore, thought worthy of further investigation. In particular, the question to address is why QALYs gained by those who would otherwise die instantly (i.e. whose QoL lost would be 1) were not weighted higher than other types of QALY gain.

To investigate this further, we isolated choice sets in which one scenario involved QoL lost being equal to 1 (i.e. age at onset = age at death), referred to here as the ‘life-saving’ option. By selecting choices with just one life-saving option, by definition, the comparator scenario within any such choice would be any of all other types (e.g. QoL enhancing only and L-shaped gains which enhance QoL, but also add years to life). This amounted to 1253 choice sets in total.

On 446 occasions, the life-saving option was chosen. On these occasions, the mean QALY gain being offered in the life-saving option was 12.8 (median = 7) as opposed to a mean QALY gain of 10.7 (median = 7.8) in the alternative (which was not chosen). The life-saving option, by definition, presents QoL lost as 1, and the mean QoL lost presented in the rejected scenario was 0.54. On these occasions it seems that, on average, QALY gains are larger and beneficiaries are more severe in life-saving options, so it is not surprising that this was chosen.

On the 809 occasions when life saving was not chosen, the mean QALY gain being offered in the life-saving option was 5.43 (median = 4) as opposed to 13.4 (median = 7.9) in the alternative; the alternative having a mean QoL lost of 0.53. On these occasions it seems that, on average, QALY gains are much larger when beneficiaries are in a less severe state. This seems to have driven the choices made, making severity appear less important than it might be if choices were offered in which the scenarios were closer together in terms of numbers of QALYs gained. This may explain why results of matching-type studies, which essentially set QALYs equal and highlight differences in equity variables of interest (such as severity), tend to reveal more enhanced weights for such variables with beneficiaries in more severe states being more highly valued. An improved design might have meant that the interaction between QALYs and severity could have been allowed for more comfortably, which may have shown that the shape of the relationship between severity and utility was positive when treatment provides large QALY gains – a preference shown for the most severe – and negative when treatment provides small QALY gains – a preference for the least severe. That is, when small amounts of QALYs are being offered to someone who is already in a very bad state (or who will die imminently) it might seem reasonable not to place additional value or weight on that more severe state. Indeed, this is one interpretation of the results from one other study we are aware of, which provides a result that runs counter to the usual view of severity;⁵⁹ however, our data are not able to show this one way or the other. Other studies on severity, based on matching-style approaches, are discussed in Chapter 5, Concluding remarks.

Concluding remarks

The discrete choice part of the relativities study has exposed the reader to three things: the challenges

of deriving relative weights for QALYs using discrete choice methods; two novel approaches to eliciting relative QALY weights from a discrete choice experiment; and some initial empirical weights.

Overall, the estimated weights are similar across different regression models and different weighting procedures. There are obviously differences between the simple model and that including higher-order polynomials due to the imposed restriction of monotonicity in the former. Comparing the CV weights with those from the probability-of-choice approach, we can see that the range of weights is smaller (0.8–1.1) for the former, and that they are generally nearer to 1, suggesting no difference between the base case and the comparison scenario. More generally, this would indicate that age and severity have little impact on QALY gains. The range for the probability-of-choice approach is 0.38–1.61. As the weights are based on the same regression model, it is unsurprising that the same scenarios result in the lowest and highest weights in both methods.

The narrow range of weights, especially using the CV approach, and the lack of statistical significance of age at onset and severity in the regression models, raise a number of issues. It might be that, when presenting information on such variables within a context in which health gains also vary, the impacts of age and severity are diminished. Indeed, this could be seen as an advantage of a discrete choice approach in which equity attributes are embedded within scenarios in which health gains (or QALYs) are also allowed to vary. However, it could also be the case that this result is driven by the compromises made necessary in the experimental design.

There are some clear patterns. For example, as the age at onset increases, the weights move towards 0 (until age 20) and then move towards 1. This shows a preference for those who fall ill between the ages of 10 and 40, with 20 being the most preferred. As age at death increases, the weights move towards 0 (until age 10) and then move towards 1 (up to age 40), before moving back towards 0. This shows a preference for individuals who will die at 1, 10 and 80 and less of a preference for those between 20 and 70. As severity increases, there is a preference for those who are slightly ill, a QoL loss of 0.4 and 0.1, compared with those who lose 0.7 and 1. Of course, the advantage with these approaches is that it is possible that age effects outweigh severity effects, or vice versa, meaning that generalisations are harder to make unless we can invoke *ceteris*

paribus assumptions. In other words, different weights on QALYs are required for a variety of contexts.

With respect to more methodological issues, we have demonstrated two novel approaches to deriving weights from our empirical data. First, what issues arise from people's perceived similarities or differences in the weights? At a conceptual level, the CV and probability-of-choice approaches to deriving relative weights are similar. However, instead of equalising the probability of choice between the two health states, the CV method equalises the utility associated with the two health states – a subtle but important difference. Second, does this difference matter? One possibility, of course, is that using the CV approach overcomes one limitation of the probability approach in that the CV can be used to calculate weights per attribute in addition to weight for entire health-state scenarios.

The controversy of these results lies in the small impact of the equity variables of interest (age and severity). However, this should not necessarily be taken as an indication that weighting of QALYs is

not desirable. In the short term, the results need to be reinforced by calculation of confidence intervals around the weights that have been presented. This is a significant task; hence such data have not been presented in this report. In addition, despite the controversial nature of this result, as mentioned above, it may have arisen due to one potential advantage of a discrete choice approach, whereby variations in equity attributes are considered alongside variations in health gains. Our main longer-term recommendation would, therefore, be for more research, as replication of such a result would be crucial prior to any subsequent policy recommendations. This research should also pursue alternative experimental design strategies in order to address the issue of implausible scenarios while still maintaining desirable design properties, so ensuring we can estimate the effects of interest with improved efficiency. Further important potential challenges, such as the nature of the (multiplicative) functional form used in the statistical analysis and the transformation of variables from the original forms in which they were presented in *Table 6*, are discussed further in Chapter 5, Concluding remarks.

Chapter 5

Matching study

Basic approach

We report here the ‘matching’ relativities study which, along with the DCE study (see Chapter 4), set out to determine whether the UK public wish to attach more weight to some QALYs than to others. To recap, a QALY is 1 year in full health and years spent in less than full health are ‘quality adjusted’ in order to take account of the morbidity associated with disability, disease or illness. As QALYs combine morbidity (QoL) and mortality (length of life) on a single scale, they allow comparisons to be made across interventions with different types of health outcomes (e.g. QoL enhancing versus life extending). In the standard ‘unweighted’ QALY model, all QALYs are of equal value. For example, the age of recipients does not matter, as long as the QALY gain is the same. Likewise, the standard model assumes that equal QALY gains are of equal value regardless of how severely ill the patients are prior to treatment. The aim of the matching – and DCE studies – is then to address the question of whether a QALY is a QALY. Or put another way, is the ‘standard’ model correct?

While each respondent in the survey answered a set of matching and discrete choice questions, we concentrate here on the results of the matching study, the main aims of which were:

- to estimate the relative weights to be attached to a QALY according to the age of recipients
- to estimate the relative weights to be attached to a QALY according to the severity of illness of recipients.

The secondary aims of the matching study were to assess the impact of certain other attributes considered to be of interest to respondents and which featured in the design such as gain in life expectancy and gain in QoL.

Methods

Matching questions

Briefly, ‘matching’ questions ask people to state the number of outcomes of one kind they consider to

be ‘just as good as’ a specified number of outcomes of another kind. A typical matching question in this setting would be something along the following lines:

Consider two groups of 100 people: 100 of type A and 100 of type B (these types being described and differentiated by age and severity of illness prior to treatment). With treatment, each of these types could experience some given health benefit. If there were only enough resources to treat one group, which would you prefer it to be?

Suppose the respondent chooses to give priority to Group A. The number of beneficiaries in Group B is held fixed at 100 while the number of A-types is reduced to the point where the respondent finds it hard to prioritise between that number (X) of A-types and 100 B-types. By systematically varying the age and health of the people described in Groups A and B in a series of questions, it is possible to investigate the relative weights attached to health gains in people of different ages and levels of health.

The QALY grid approach

Recall that the animated POWERPOINT introduction to the CAPI used a graph with QoL (from 0% to 100%) on the vertical axis and Age (from 0 to 80) on the horizontal axis. Clearly, all questions posed in the relativities studies – both matching and DCE – must involve health gains that fall within this defined space. Further, all questions must be ‘logical’ in that there must be some positive gain in terms of either QoL or length of life, or both. While it was necessary to incorporate such considerations explicitly into the DCE design (the underlying principle of which is to allow all attributes to vary independently), ‘illogical’ scenarios were ruled out a priori in the matching design. There remain, however, an enormous number of potential pairings that can be made, and the challenge was to come up with the most parsimonious design that allowed all the research questions to be addressed. We were keen that the design be capable of testing the standard ‘unweighted’ QALY model, i.e. where all relative

weights would be equal to one, as well as being able to detect systematic departures from this baseline assumption with respect to the age of patients and their ‘untreated’ QoL. To that end we developed the ‘QALY grid’ approach, summarised briefly below.

Consider again the graph with QoL (from 0% to 100%) on the vertical axis and Age (from 0 to 80) on the horizontal axis. By partitioning the vertical into five 20% ranges and dividing the horizontal into four 20-year intervals, we get 20 ‘cells’ (numbered 1–20 in *Table 13*).

Improving a person’s health by any one cell, i.e. by 20% per year for 20 years, would, under the ‘standard’ model, give four QALYs. But are all QALYs weighted equally, irrespective of where they are located in the grid? For example, suppose that type A people were those whose QoL would, if untreated, be 60% at (and after) age 20 and treatment could raise that QoL to 80% for the next 20 years, whereas type B people would be at 80% from the age of 60 but could receive treatment that would restore them to 100% for the subsequent 20 years. Both types would gain four ‘standard’ QALYs; but do respondents give both treatments equal priority? If not, which group of 100 do they give priority to? And if the number of beneficiaries in that group is reduced, at what value of $X < 100$ do the two treatments receive equal priority? For example, if a respondent finds it hard to prioritise when the same total expenditure of resources could treat either 40 people of type A or 100 people of type B, that would suggest that such a respondent gives two-and-a-half times as much weight to a QALY in cell 14 as in cell 20.

If we have data on respondents’ relative values for the 20 cells, and the value of a ‘reference’ cell is set equal to 1 (suppose we call that v_1), we can derive

a set of relative weights (v_2 – v_{20}) for the remaining 19 cells in the grid. Our basic aim was to estimate these relative values. If each cell is paired with every adjacent cell, we can estimate ratios such as v_1/v_2 and v_1/v_3 . This requires 31 pairings. An example of a question respondents were actually presented with in the matching study is reproduced in *Figure 14*.

While this is a parsimonious design, extrapolating from these 31 ‘single cell’ questions requires a number of strong assumptions about the nature of respondents’ preferences; assumptions we did not feel able to make without first testing their validity.

So another 41 questions in the overall design are intended to embody a number of such tests. Some of these questions look at blocks of cells: for example, each row P–T is paired with every adjacent row, and each column with every adjacent column. One important difference between the ‘whole column’ and ‘single cell’ pairings is that the former *necessarily* involve ‘life-extending’ QALYs (as they offer an additional 20 years of life expectancy). For example, pairing cells 10 and 11 asks respondents to consider an additional four QALYs (moving from 40% to 60% health for 20 years) either between the ages of 20 and 40 or between the ages of 40 and 60. In contrast, pairing columns B and C asks respondents to consider extending life (at 100% health) for 20 years at the age of either 20 or 40. If respondents were to have preferences over ‘life-extending’, rather than ‘life-enhancing’ QALYs, we might expect the results of the questions involving the whole column to be somewhat different than implied by the ‘sum of the parts’, i.e. those pairings of the single cells that make up the columns.

The design also allows us to test for consistency via chaining with equal numbers of steps, e.g.

TABLE 13 The QALY grid

QoL	A	B	C	D	
80–100%	17	18	19	20	P
60–80%	13	14	15	16	Q
40–60%	9	10	11	12	R
20–40%	5	6	7	8	S
0–20%	1	2	3	4	T
Age	0–20	20–40	40–60	60–80	
QoL, quality of life					

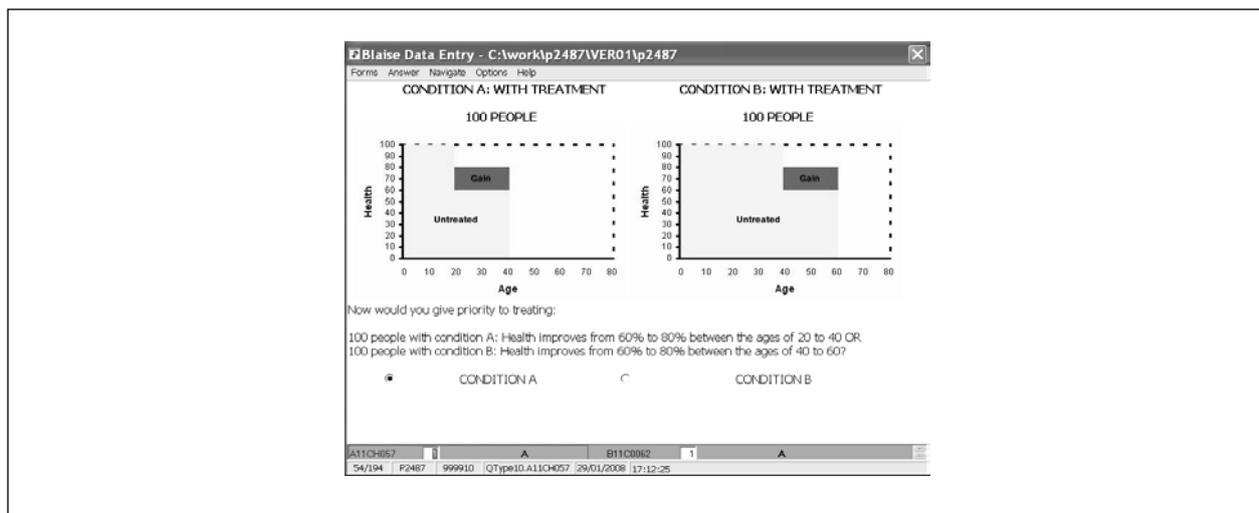


FIGURE 14 Matching question from CAPI.

$(v_1/v_3) \cdot (v_5/v_9)$ should equal (v_1/v_9) . To test for chaining with different numbers of steps, we have built in some *non-adjacent pairings*. This has been done for *single cells* (e.g. compare the pairs {1, 2}, {2, 3} and {1, 3}), for *whole columns*, e.g. compare {A, B}, {B, C} and {A, C} and for *whole rows*, e.g. compare {R, S}, {S, T} and {R, T}.

Finally, there are the tests for independence of the profile in which a cell appears. Consider the different ways in which cell 14 may appear in a scenario. In *Figure 15*, 14S shows the patient to be in good health until aged 20 when – without treatment – their QoL falls from 100% to 60% and they die at the age of 40.

This depicts the ‘standard’ representation that applies to all other single cell comparisons used here. In 14X the patient is at 80% of good health until the age of 20, when their QoL falls to 60%, and again they die at the age of 40. In 14Z, the patient is at 80% of good health until the age of 20, when their QoL falls to 60%, but now they live until the age of 80. Finally, in 14Y the patient is in good health from birth and lives to the age of 80. Each profile test compares the ‘standard’ representation with one of the three other representations of that cell (i.e. 14S versus 14X, 14S versus 14Z, 14S versus 14Y). Three different sets of profile tests were carried out, involving cells 7, 11 and 14. ‘Fair innings’ type arguments would suggest a tendency, all else being equal, to favour those patients whose total lifetime QALYs were lower.

For each of the three cells used in the profile tests, a ‘fair innings’ type argument would predict that

more weight would be given to S than to Y (as the total lifetime QALYs is less in S than in Y), but S would be given less weight than X (as the total lifetime QALYs is greater in S than in X). Although not shown in detail here, S would be given less weight than Z in the cases of cells 7 and 11, but S would be given more weight than Z in the case of cell 14.

The 72 questions were spread evenly across 12 versions of the matching questionnaire, each respondent answering six matching questions (along with eight DCEs).

Aggregating results

Suppose the following represents data from five respondents who were indifferent between treating X patients in Group A and Y patients in Group B (with either X or Y = 100):

	Group A	Group B
	X	Y
Resp 1	100	50
Resp 2	100	90
Resp 3	100	50
Resp 4	50	100
Resp 5	90	100

This pattern indicates that respondents 1 to 3 preferred to give priority to Group B at the first matching iteration, i.e. when there were 100 in each group, and subsequently set the number treated in that group (Y) at some number lower

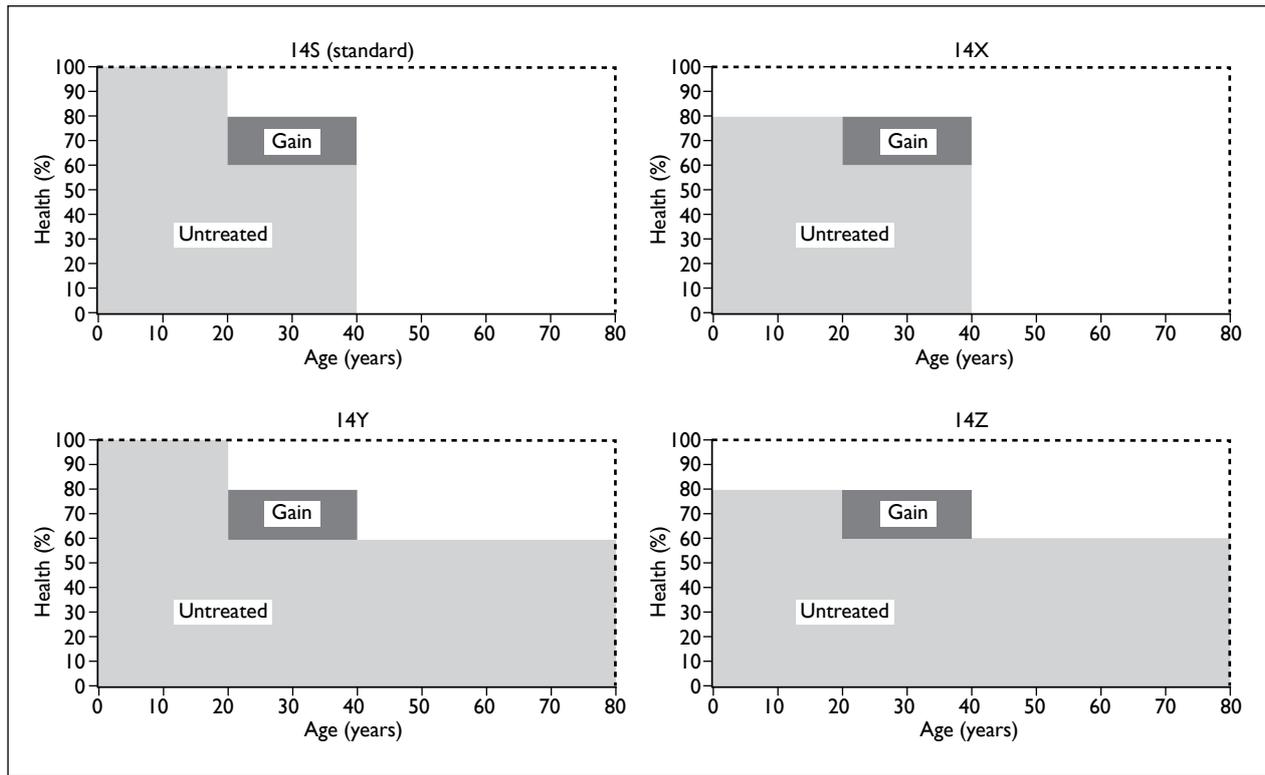


FIGURE 15 The profile tests (cell 14).

than 100. In contrast, respondents 4 and 5 preferred to give priority to Group A initially, and subsequently set the number treated in that group (X) at some number lower than 100. Clearly, respondents 3 and 4 have identical, but opposite, preferences. Yet, the ratio X/Y is 2:1 for respondent 3 and 0.5:1 for respondent 4 (and vice versa for Y/X) resulting in a very different effect on the arithmetic mean of individual ratios. As the decision to take the ratio X/Y or Y/X is arbitrary, this asymmetry is an undesirable property of an aggregation method.

There are two ways in which we shall aggregate the results of the matching questions in order to overcome problems due to lack of symmetry. The first of these – the ratio of means method – is the method used previously in work commissioned for the Health and Safety Executive (HSE), which set out to investigate whether premia ought to be attached to the prevention of certain types of deaths.⁶⁰

The ratio of means method

In this method the most favoured group attracts a value of 1 and the less favoured a value equal

to the number of patients treated in the more preferred group divided by the number treated in the less preferred group. For example, Group B was respondent 1’s ‘most preferred’ group and they set Y at 50. Group B then attracts a value of 1 and Group A a value equal to $Y/X = 0.5$. Following this principle for all five respondents yields the following values:

	Group A	Group B
Resp 1	0.5	1
Resp 2	0.9	1
Resp 3	0.5	1
Resp 4	1	0.5
Resp 5	1	0.9

Means of each column are then taken and the ratio of those means are calculated. In this case, the mean value for Group A is 0.78 while the mean value for Group B is 0.88. The implied weight of Group B relative to Group A is then $0.88/0.78 = 1.13$. Alternatively, the implied weight of Group A relative to Group B is $1/1.13 = 0.88$, i.e. the measure is symmetrical.

Aggregating ratios derived from individual responses

Alternatively, the ratios X/Y may be computed for each individual respondent and then the median can be identified. The ratios are as follows:

	X/Y
Resp 1	2:1
Resp 2	1.11:1
Resp 3	2:1
Resp 4	0.5:1
Resp 5	0.9:1

Taking medians of the individual ratios results in a weight of Group A relative to Group B of approximately 1.11. As with the ratio of means method above, this measure is symmetrical, i.e. the median of the Y/X ratios is $1/1.11 = 0.9$.

In this particular example, both measures produce very similar measures of central tendency. However, as we shall see below, this may not always be the case.

Results

Responses to first matching question

Before going on to look at the weights derived from the aggregate matching data, we begin by looking at the number of respondents choosing to give priority to one group or the other in the first matching iteration, i.e. when the number of patients treated was 100 in both groups. These data are the most straightforward to interpret and are arguably the simplest way of illustrating the *general patterns* to emerge from the relativities study.

With regard to the general pattern relating to age, a total of 269 respondents answered a matching question involving an equivalent gain going either to 40- to 60-year-olds or to 60- to 80-year-olds. Of these 269, 173 (64%) preferred to give priority to the younger patients. Similarly, a total of 294 respondents answered a matching question involving an equivalent gain going either to 20- to 40-year-olds or to 40- to 60-year-olds. Of the 294, 182 (62%) preferred to give priority to the younger patients. In contrast, a total of 295 respondents answered a matching question involving an equivalent gain going either to 0- to 20-year-olds or to 20- to 40-year-olds. Of the 295, only 107 (36%) preferred to give priority to the younger

patients. This suggested that the pattern relating to age is not a simple one and that 'non-linearities' may exist with respect to age weights.

This pattern is borne out if we look only at responses to the whole column questions (detailed above) which involved gains of 20 QALYs that were necessarily life extending. Of the 51 respondents who were asked to compare column A (20 years in full health for a newborn) with column B (20 years in full health for a 20-year-old), 17 (33.3%) preferred to treat the younger patients. In contrast, of the 54 respondents who compared column C (20 years in full health for 40-year-olds) with column D (20 years in full health for 60-year-olds), 36 (66.7%) preferred to treat the younger patients.

With regard to the general pattern relating to severity, a total of 294 respondents answered a matching question involving a move either from 60% to 80% health or from 80% to 100% health (for a given age group). Of the 294, 170 (58%) preferred to give priority to patients in the more severe health state. In contrast, of the 280 respondents who answered a matching question involving a move either from 0% to 20% health or from 20% to 40% health (for a given age group), only 105 (38%) preferred to give priority to those patients in the more severe health state. Again, observing these initial choices suggested that non-linearities may exist with respect to severity weights. The pattern is borne out by observing responses to the 'whole row' questions. Of the 44 respondents who were asked to compare row T (a move from 0% health to 20% health between birth and age 80) and row S (a move from 20% health to 40% health between birth and aged 80), 17 (38.6%) preferred to treat those patients who were worse off initially. In contrast, of the 44 respondents who were asked to compare row Q and row P, 32 (71.1%) preferred to treat those patients who were worse off initially.

As it seemed plausible that the desire to prioritise one age group over another may be related to the age of respondents, we also broke these data down by age of respondent. The sample was divided into three broad age groups: the under 40-year-olds, the 40- to 60-year-olds and the over 60-year-olds. These categories were chosen as they best coincided with those used in the QALY grid (as respondents had to be 18 or over to participate in the survey, there are too few respondents to have a separate category for 0- to 20-year-olds). In *Tables 14* and *15*, the numbers in the cells relate to those

TABLE 14 Preferences for treating 0- to 20-year-olds over 20- to 40-year-olds by respondent age

Age of respondent	Preference given to 0- to 20-year-olds	Preference given to 20- to 40-year-olds	n
Under 40	48 (49%)	51 (51%)	99
40–60	35 (34%)	67 (66%)	102
Over 60	24 (26%)	70 (74%)	94
Total	107	188	295

$\chi^2(2) = 11.79; p = 0.003.$

TABLE 15 Preferences for treating 40- to 60-year-olds over 60- to 80-year-olds by respondent age

Age of respondent	Preference given to 40- to 60-year-olds	Preference given to 60- to 80-year-olds	n
Under 40	60 (68%)	21 (32%)	81
40–60	65 (61%)	30 (39%)	95
Over 60	48 (49%)	45 (51%)	93
Total	173	96	269

$\chi^2(2) = 7.65; p = 0.022.$

respondents preferring to treat one group over the other at the initial iteration in the matching question, i.e. when there were 100 patients in both groups. The numbers in brackets relate to the percentage of respondents in each age group preferring to treat that group over the other. In each case, differences across age groups have been tested using chi-squared tests with two degrees of freedom.

Thus, it does appear as if responses are significantly different according to age of respondent. The general pattern to emerge from the data (including some not shown here) is that older respondents seem more inclined than younger respondents to give preference to older patients, but this does not appear to be related just to a self-interested desire to prioritise their own age group (as they are also more likely to give preference to 20- to 40-year-olds than to 0- to 20-year-olds).

As it also seemed plausible that a desire to prioritise in terms of severity may also be related to respondent age, we broke down the data relating to severity in this way. In *Tables 16* and *17* the numbers in the cells represent respondents preferring to treat one group over the other, now defined by severity of health state (and holding age of patients constant).

Again, it does appear as if responses are significantly different according to age of respondent. In particular, *Table 17* shows that older respondents are more likely to favour that group in better health initially when the gains take place towards the top of the severity spectrum. This may reflect the fact that older respondents place greater weight on returning patients to full health than their younger counterparts.

While observing that these first choices is a useful means of illustrating the *general* patterns to emerge from the data, we go on to look at the aggregate results from the iterative matching procedures that followed.

The relative weights

Recall that our ‘QALY grid’ partitions the vertical axis of *Figure 14* into five 20% ranges and the horizontal axis into four 20-year intervals, giving 20 ‘cells’, each worth four QALYs under the ‘standard’ model. We then set out to determine whether all QALYs were weighted equally, irrespective of where they are located in the grid. We begin by presenting the results of the 31 single cell comparisons and the ratio of means aggregation method. The implications of using alternative aggregation methods are outlined below.

TABLE 16 Preferences given to gain from 0% to 20% health over gain from 20% to 40% health (for given age of patient) by age of respondent

Age of respondent	Preference given to gain from 0% to 20%	Preference given to gain from 20% to 40%	n
Under 40	38 (46%)	45 (54%)	83
40–60	28 (28%)	73 (72%)	101
Over 60	39 (41%)	57 (59%)	96
Total	105	175	280

$\chi^2(2)=7.30; p=0.026.$

TABLE 17 Preferences given to gain from 60% to 80% health over gain from 80% to 100% health (for given age of patient) by age of respondent

Age of respondent	Preference given to gain from 60% to 80%	Preference given to gain from 80% to 100%	n
Under 40	43 (68%)	20 (32%)	63
40–60	75 (61%)	49 (39%)	124
Over 60	52 (49%)	55 (51%)	107
Total	170	124	294

$\chi^2(2)=7.30; p=0.026.$

TABLE 18 Horizontal (age) weights: ratio of means method (cell numbers are shown in square brackets)

QoL				
80–100%	[17] 0.837	[18] 1.559	[19] 1.834	[20] N/A
60–80%	[13] 0.781	[14] 1.131	[15] 1.564	[16] N/A
40–60%	[9] 0.676	[10] 1.039	[11] 1.383	[12] N/A
20–40%	[5] 0.716	[6] 1.298	[7] 1.433	[8] N/A
0–20%	[1] 0.867	[2] 1.228	[3] 1.687	[4] N/A
Age	0–20	20–40	40–60	60–80

N/A, not applicable; QoL, quality of life.

The numbers in each cell of *Table 18* signify the weight attached to that cell relative to the cell to the right, as derived via the method set out under 'Aggregating results'. Values greater than 1 indicate that more weight is attached to that cell than to the one to the right. Values less than 1 indicate that less weight is attached to that cell than to the one to the right. For example, a value of 1.564 in cell 15 indicates that a gain (in this case from 60% to 80% health) accruing to 40- to 60-year-olds is valued at 1.564 times the equivalent gain accruing to 60- to 80-year-olds. In contrast, the value of 0.676 in cell 9 indicates that a gain (in this case from 40% to 60% health) accruing to 0- to 20-year-olds is worth

0.676 times the equivalent gain accruing to 20- to 40-year-olds.

Following this through for all other adjacent horizontal cell comparisons, the general pattern to emerge with respect to age (holding severity constant) is as follows:

- Less weight is given to treating 0- to 20-year-olds than 20- to 40-year-olds.
- More weight is given to treating 20- to 40-year-olds than 40- to 60-year-olds.
- More weight is given to treating 40- to 60-year-olds than 60- to 80-year-olds.

TABLE 19 Vertical (severity) weights: ratio of means method (cell numbers are shown in square brackets)

QoL				
80–100%	[17] N/A	[18] N/A	[19] N/A	[20] N/A
60–80%	[13] 1.064	[14] 1.369	[15] 1.203	[16] 1.618
40–60%	[9] 0.792	[10] 1.144	[11] 1.179	[12] 1.247
20–40%	[5] 0.955	[6] 1.185	[7] 1.235	[8] 1.247
0–20%	[1] 0.786	[2] 0.759	[3] 0.857	[4] 0.801
Age	0–20	20–40	40–60	60–80
N/A, not applicable; QoL, quality of life.				

Though not shown in detail here, this broad pattern is borne out in the ‘whole column’ comparisons, indicating that this pattern holds for life-extending QALYs involving larger gains too.

Turning now to the vertical – or severity – weights, the numbers in each cell of *Table 19* signify the weight attached to that cell relative to the cell above it. Values greater than 1 indicate that more weight is attached to that cell than to the one above it and vice versa. For example, a value of 1.618 in cell 16 indicates that, for a given patient age group, a gain from 60% health to 80% health is valued at 1.618 times that from 80% to 100% health. In contrast, a value of 0.801 in cell 4 indicates that a gain from 0% to 20% health is valued at 0.801 times that from 20% to 40% health (again keeping age of patient constant).

The general pattern to emerge with respect to severity (holding age constant) is as follows:

- Less weight is given to a move from 0% to 20% than from 20% to 40%.
- More weight is given to a move from 20% to 40% than from 40% to 60%.

- More weight is given to a move from 40% to 60% than from 60% to 80%.
- More weight is given to a move from 60% to 80% than from 80% to 100%.

(The exception to this pattern is the 0–20 age range, where a move from 60% to 80% is weighted most highly and where the weights drop away progressively above and below that cell.)

Given the broad regularity of the patterns for both age and severity – in both cases, the general shape is a (somewhat off-centre) inverted U – one simple way of combining the two into a single set of weights for the 20 cells is as follows:

1. Identify the average pattern showing how weights vary just with age, independent of severity.
2. Likewise, identify the average pattern showing how weights vary just with severity, independent of age.
3. Compute the weight for any cell as the cross-product of the appropriate age and severity weights, normalising overall so that the most highly weighted cell is indexed at 1.

TABLE 20 Horizontal (age) weights inferred from Table 18 (cell numbers are shown in square brackets)

QoL				
80–100%	[17] 0.837	[18] 1.000	[19] 0.641	[20] 0.350
60–80%	[13] 0.781	[14] 1.000	[15] 0.884	[16] 0.565
40–60%	[9] 0.676	[10] 1.000	[11] 0.962	[12] 0.697
20–40%	[5] 0.716	[6] 1.000	[7] 0.770	[8] 0.538
0–20%	[1] 0.867	[2] 1.000	[3] 0.814	[4] 0.483
Age	0–20	20–40	40–60	60–80
Average for each age range	0.775	1.000	0.814	0.527
QoL, quality of life.				

One way of achieving that result is as follows. It is clear from *Table 18* that at every level of severity the 20–40 age range is most highly weighted, so assign those cells an ‘age weight’ of 1 and use the ratios reported in *Table 18* to infer relative weights for the other cells on each row. The result is reported in *Table 20*. The average weights for each age range are shown in the bottom row.

Likewise, noting that the highest weight on the severity axis for three of the four age ranges is accorded to a move from 20% to 40%, we can

assign weights of 1 to those cells and infer relative weights for the other cells in each column (*Table 21*). Then the average for each severity level can be calculated; this is shown in the last column.

Taking the cross-product of the average weights for age and severity gives *Table 22*.

Table 22 gives the estimated weight of each cell relative to the cell with the highest value – cell 6, where a gain from 20% to 40% health accrues to 20- to 40-year-olds. The general pattern is that the

TABLE 21 Vertical (severity) weights inferred from *Table 19* (cell numbers are shown in square brackets)

QoL						Average for each severity
80–100%	[17] 1.243	[18] 0.539	[19] 0.571	[20] 0.397		0.688
60–80%	[13] 1.322	[14] 0.738	[15] 0.687	[16] 0.643		0.848
40–60%	[9] 1.047	[10] 0.844	[11] 0.810	[12] 0.802		0.876
20–40%	[5] 1.000	[6] 1.000	[7] 1.000	[8] 1.000		1.000
0–20%	[1] 0.786	[2] 0.759	[3] 0.857	[4] 0.801		0.791
Age	0–20	20–40	40–60	60–80		
QoL, quality of life.						

TABLE 22 Implied relative weights: ratio of means method (cell numbers are shown in square brackets)

QoL					
80–100%	[17] 0.533	[18] 0.688	[19] 0.560	[20] 0.362	
60–80%	[13] 0.658	[14] 0.848	[15] 0.690	[16] 0.446	
40–60%	[9] 0.679	[10] 0.876	[11] 0.713	[12] 0.461	
20–40%	[5] 0.775	[6] 1.000	[7] 0.844	[8] 0.527	
0–20%	[1] 0.613	[2] 0.791	[3] 0.644	[4] 0.417	
Age	0–20	20–40	40–60	60–80	
QoL, quality of life.					

TABLE 23 Implied relative weights: median of individual ratios method (cell numbers are shown in square brackets)

QoL					
80–100%	[17] 0.47	[18] 0.72	[19] 0.54	[20] 0.24	
60–80%	[13] 0.55	[14] 0.85	[15] 0.64	[16] 0.29	
40–60%	[9] 0.60	[10] 0.92	[11] 0.69	[12] 0.31	
20–40%	[5] 0.65	[6] 1.00	[7] 0.75	[8] 0.34	
0–20%	[1] 0.49	[2] 0.75	[3] 0.56	[4] 0.26	
Age	0–20	20–40	40–60	60–80	
QoL, quality of life.					

valuation function peaks at cell 6, but then ‘falls away’ in all directions. As the top right-hand corner cell combines the lowest average weight assigned to severity with the lowest average weight assigned to age, it carries the lowest weight of any cell; on this basis, the differential between the highest and lowest valued cells is roughly 2.75:1.

As noted under ‘Aggregating results’, the ratio of means method is one method of aggregating the matching data. However, another set of weights can be produced on the basis of medians. The method of construction is essentially the same: the median ratios between rows and between columns allow us to fix the weight of cell 6 at 1, and set the weights for all other cells in the same row and column as cell 6, on the basis of those medians; the weights in all other cells are then computed as the appropriate cross-products. The results are given in *Table 23*.

Not surprisingly, weights based on means do not exactly coincide with the median-based measures. The averaging process appears to make the slope away from the ‘peak’ weight less steep in nearly all directions, so that the median-based weights are somewhat lower in most cases: while the most extreme comparison from *Table 22* is 2.75:1, the corresponding ratio in *Table 23* is approximately 4:1. However, although the specific weights may vary according to the measure of central tendency chosen and the particular method of aggregation, the general pattern of weights dropping as we move in any direction away from cell 6 is robust to these differences.

Chaining tests

As outlined in the methods sections, a number of non-adjacent pairings were carried out in order to test for chaining with different numbers of steps with the basic idea being that $(v_1/v_5)*(v_5/v_9)$ should roughly equal (v_1/v_9) . In *Table 24*, the

TABLE 24 Results of ‘chaining’ tests

	Direct	Chained
9 vs 17	0.92	0.84
11 vs 19	1.18	1.42
10 vs 12	1.38	1.44
18 vs 20	2.12	2.85
6 vs 8	1.83	1.86
A vs C	0.93	0.71
R vs P	1.91	2.46

‘direct’ column reports actual responses to the non-adjacent pairings, i.e. (v_1/v_9) in the example used here, while the values in the ‘chained’ column are derived by multiplying together the weights of the component parts of the ‘chain’, i.e. $(v_1/v_5)*(v_5/v_9)$ in the example used here. Chaining tests were also carried out for whole columns, e.g. compare {A, B}, {B, C} and {A, C} and for whole rows, e.g. compare {R, Q}, {Q, P} and {R, P}. In each case the values relate to the aggregate matching results and make use of the ratio of means aggregation method.

For each of the non-adjacent pairings in the chaining tests we can see that the value implied by ‘chaining’ is more extreme than that derived directly. For example, the second row of *Table 24* shows that $(v_{11}/v_{15})*(v_{15}/v_{19}) = 1.42$ while $(v_{11}/v_{19}) = 1.18$. Likewise, row 4 shows that $(v_{18}/v_{19})*(v_{19}/v_{20}) = 2.85$ while $(v_{18}/v_{20}) = 2.12$.

Profile tests

Finally, we turn to the tests for independence of the profile in which a cell appears. Recall that *Figure 15* illustrated the alternative ways in which we might have presented the QALY gains embodied in the cells in the QALY grid. Profile tests were carried out on three different cells: 7, 11 and 14. We have argued that a ‘fair innings’ type argument – in which the total lifetime QALYs was relevant – would always predict that *more* weight would be given to the ‘standard’ profile (S) than to profile Y (as the total lifetime QALYs is less in S than in Y), and *less* weight to profile S than profile X. Finally, profile S would be given less weight than Z in the cases of cells 7 and 11, but S would be given more weight than Z in the case of cell 14.

The results are given in *Table 25* where values greater than 1 indicate that the gain was valued more highly when embedded in the profile S than in the alternative and vice versa. For example, the first column of *Table 25* gives the results of comparing the profile S with profile X (see *Figure 15*). As the values are each greater than 1, there appears, if anything, to be a tendency to favour

TABLE 25 Results of the ‘profile’ tests

	S vs X	S vs Y	S vs Z
Cell 7	1.33	0.74	1.24
Cell 11	1.22	1.36	1.08
Cell 14	1.06	0.55	0.78

those with higher lifetime health, which is the opposite of what a 'fair innings' argument would suggest. Column 2 shows that for two of the three cells (7 and 14) the standard profile was valued less highly than profile Y, which had lower lifetime health – a finding which does fit with a 'fair innings' type argument – but that the opposite was true in the case of cell 11. The results of the S versus Z comparison also appear to contradict the 'fair innings' argument.

To sum up, it is difficult to detect any clear pattern in responses to the profile tests conducted here, but, if anything, the data appear to contradict the 'fair innings' argument, which is puzzling.

Concluding remarks

We reported here the results of a computer-assisted 'matching', or person trade-off, valuation study carried out in a sample of 587 members of the public in the UK. The results showed that, when asked questions of the type posed here, respondents *do* differentiate between 'types' of QALYs and are willing to trade off (sometimes considerable) numbers of patients treated in order to prioritise according to age and severity of illness. The matching data show a general tendency to give more weight to younger patients and to those in poorer health, but the pattern is not a simple one. In particular, less weight is given to the youngest patients and to those in poorest health, reversing the pattern found elsewhere in the data. As these patterns may be observed in the simple choice data from the first matching iteration (when there were 100 patients in each group), they are robust to differences in methods used to aggregate the data.

The qualitative findings reported in Chapter 3 may offer some explanation for the pattern of results uncovered here. Recall that age was found to be important to focus group participants for a variety of reasons including life expectancy, which would explain the general tendency to favour the young over the old. Recall also, however, that age was found to be used as a 'proxy' for a range of other characteristics including economic activity and the existence of dependants. One possible explanation for the tendency to favour 20- to 40-year-olds over younger patients, therefore, is that such considerations outweigh the impact of life expectancy around those ages. If we were to represent the 20 discrete choice weights with the 20 cells of the QALY grid, the peak (most preferred cell) would be cell 14. This is consistent with the

most preferred age group from the matching data (20–40), but implies that less-severe states are preferred (60% to 80%), although the magnitude of the weights is obviously very different.

Returning specifically to the matching data, it is also reported in Chapter 3 that respondents expressed the view that an improvement in QoL for people in poor health would be more important than an identical improvement in QoL for people in relatively good health. Again, this would seem to be consistent with the general tendency to favour those in more severe health states initially. Some concern was also expressed that prolonging life in a really bad health state was not desirable. One possible explanation for why respondents favoured a move from 20% to 40% health over one from 0% to 20% is that 20% was still considered to be not worth living. Indeed, one of the respondents quoted in Chapter 3 almost makes this point explicitly: 'whereas 20% is pretty close to death.' This raises issues about how the QoL axis – which necessarily relied on a 'visual analogue scale (VAS)-like' representation in order to put the concept across to respondents – was being interpreted, particularly towards the lower end.

It is also possible that the tendency for the weights to 'dip' towards the top right-hand corner of the grid may be linked to how respondents were interpreting the QoL scale. While the diagrammatic representation of QoL over the patient's lifetime made it clear that the QoL related to the relevant stage of the patient's life, such that 100% health at the age of 70 meant in good health *for a 70-year-old*, it may be that certain respondents were considering QoL as an *absolute* measure. If this were the case, then it might have been considered that 80% health, for example, was the best that old people could reasonably expect to achieve and little was to be achieved by trying to improve their health beyond this. If there were a certain degree of this going on when respondents were answering the matching questions, this raises the possibility that the weights estimated here are perhaps penalising the 'healthy old' rather too heavily.

While the general pattern of weights estimated for the QALY grid is robust to the different methods of aggregating the data, clearly the magnitude of the weights differs markedly. The ratio of means method of aggregating the data is the more conservative of the two methods deployed here and has been used to derive relative weights elsewhere in the economy.⁶⁰ It is important to acknowledge, however, that theoretical arguments exist in favour

of other methods of aggregating the data. It is not our intention to go into these issues further here as decisions regarding the appropriate magnitude of any weights used for policy purposes will clearly be guided by additional considerations, such as acceptability and practicality.

We conclude this section by comparing how the magnitude of the weights estimated here compares with those reported in previous studies. The review paper by Dolan and colleagues¹⁶ cites a number of studies that have set out to estimate age-weights for life-years gained that have come up with a range of values. For example, Johannesson and Johannesson⁶¹ report weights for life-years gained at ages 30, 50 and 70 years of 1.0, 0.22 and 0.1 respectively, indicating a 10:1 differential between the most preferred and least preferred age groups. Adjusting for QoL affected these estimates somewhat, with three QALYs gained for a 50-year-old or nine QALYs gained for a 70-year-old judged equivalent to one QALY gained for a 30-year-old. Similarly, Nord and colleagues¹¹ derived weights for life-years gained at ages 10, 20, 60 and 80 of 1.1, 1.0, 0.4 and 0.1 respectively, again indicating a more than 10:1 differential between the most preferred and least preferred age groups. This pattern was very similar when life-improving, rather than life-extending, interventions were considered. Finally, Busschbach and colleagues⁶²

compared QoL improvements at ages 5, 10, 35, 60 and 70 and estimated weights for the utility of health at these ages to be 0.2, 1.5, 1.0, 0.7 and 0.7 respectively, indicating a 7.5:1 differential between the most preferred and least preferred age groups. It is interesting to note that these weights are more extreme than even the less conservative of our two methods of aggregating the data.

In addition, a number of studies have looked at the impact of severity of health state and found it to be an important factor, although the results are mixed. Nord and colleagues¹¹ and Ubel and colleagues⁶³ asked subjects to compare improvements on a disability scale (where lower numbers indicate better functioning) with approximately equal distances between the levels. Respondents were asked to indicate how many patients moving from level 5 to level 1 was equivalent to moving a smaller number of patients from level 6 to level 4. Both studies reported a marked preference for treating the more severely ill patients. In contrast, a recent study by Dolan and Tsuchiya⁶⁴ found that respondents gave consistently higher priority to patients with better prospects without treatment, indicating that the more severely ill attract lower priority than the better off, a finding which remains unexplained and is perhaps more in line with the result on severity in Chapter 4.

Chapter 6

Exploring the feasibility of eliciting a monetary value of a QALY

Basic strategy

Other Government departments/agencies, e.g. the Department for Transport (DfT), the Department for Environment, Food and Rural Affairs (DEFRA) and the HSE, have drawn on the stated preferences of representative cross-sections of the population to assign monetary values to the prevention of injury, illness and premature death. If it were possible to obtain a comparable preference-based monetary value for a QALY, this would offer the prospect of being able to apply cost-benefit analysis to an even broader range of public health and safety policies in a more coherent and consistent manner – and in a way which better reflects the values of the people who are paying for, and benefiting from, those policies.

So the objective of this part of the study was to explore the feasibility of eliciting such a value and, in the process, raise and address the theoretical and practical issues involved in trying to do so.

As a starting point, consider the way that preference-based monetary values have been estimated for use in road safety appraisal. Typically, surveys have sought to elicit the maximum amounts that respondents would individually be willing to pay for (usually small) reductions in their own (and possibly others') risk of being killed or injured in road accidents. These amounts are then aggregated across individuals to arrive at an overall value for the safety improvement concerned. The resultant figure thus aims to show what the safety improvement is 'worth' to the affected group, relative to the alternative ways in which that group might have spent its money.

For example, suppose that the members of a representative sample of the population were each asked for their WTP to reduce their risk of death on the roads during the next year by 1 in 100,000. If the mean response were £12, the implication would be that 100,000 members of the population would, between them, be willing to pay a total of £1.2M to achieve an overall reduction in risk that would, on average, reduce the number of fatalities on the roads by 1. Hence, the preference-based

component of the value of preventing a (statistical) fatality (VPF) would be set at £1.2M. [While this is the largest element in the overall VPF, it is not the only one: it is usually supplemented by figures which represent the loss of net consumption entailed by premature death and some estimate of the costs (medical, police, etc.) of dealing with the consequences of a typical fatal accident.]

A similar approach could be envisaged for estimating a societal WTP-based monetary value of a QALY. The simplest case would be if people had a clear idea of what a QALY is, and could then be asked for their WTP to reduce the risk of suffering the loss of 1 QALY by 1 in 1000 (for example). If a representative sample stated a mean WTP of, say, £18 for such a risk reduction, the implication would be that 1000 members of the population would, between them, be willing to pay a total of £18,000 for some intervention that would, on average, prevent the loss of one QALY.

Other closely-related variants are possible. For example, each respondent could be asked to imagine that they face a 1 in 1000 risk of an illness which, if untreated (or treated conventionally), would entail the loss of one QALY. They could then be offered an 'insurance policy' which would ensure that they received treatment (or treatment over and above the conventional provision) that would restore them to their current/normal health and avoid that QALY loss. Again, if the average WTP for this insurance were £18, the aggregate value of the treatment that generates a one-QALY benefit would be £18,000.

That is the basic principle. In the remainder of this part of the report, we describe how we explored different ways of trying to put that principle into practice, and we discuss the issues raised both by the process and by the responses elicited.

Overview of the questionnaire

In practice, very few members of the public know what a QALY is. One possibility might have been

to begin any interview by trying to explain the concept in a general way (as was attempted in the ‘relativities’ part of the project) and then ask individuals to think about how they would envisage personal QALY gains or losses and how much they would be willing to pay to achieve such gains or avoid such losses.

We do not rule this out as one possible way of proceeding; indeed, in the ‘relativities’ study there appeared to be a reasonable degree of receptiveness for the initial broad explanation of what QALYs are. However, such an explanation would have required quite a lot of time – especially if supplemented by some attempt to get respondents to relate the concept to their own present and future lives – and our focus was upon testing some of the assumptions underlying QALYs and their possible monetisation, rather than accepting them as uncontested. So instead of asking people to value (chances of) QALY gains/losses directly, we adopted the strategy of asking people to value avoidance of, or reductions in the risk of, illness states described quite naturalistically; and we then tried to elicit in a standard way some QoL index measure of those states. Putting the two together would – if certain assumptions hold – enable us to assign money values to QALYs; but separating the two tasks would also allow us to examine more closely whether the necessary assumptions do in fact hold.

To be more specific, the questionnaire centred around two illness states – one involving recurrent stomach/bowel problems, the other involving recurrent episodes of head pain. For each state, there were three possible durations: 3 months, 12 months and ‘the rest of your life’. And for various combinations of these, there were questions about how much the respondent would be willing to pay to prevent the certainty of the illness and how much they would be willing to pay to eliminate some risk of the illness. Having obtained these monetary responses, the questionnaire went on to elicit information about trade-offs the respondent would make in terms of the relative probabilities of different health states, including some measure of the QoL index assigned to the illness state. In theory, this trade-off information makes it possible to compute for each individual the QALY losses entailed by the various illness descriptions; and this information, combined with the individual’s WTP to avoid/reduce the risk of the various illness scenarios, should (if the underlying assumptions hold) enable us to take readings of that individual’s monetary value of a QALY.

The questionnaire can be viewed as divided into four sections. Part A asked questions about each individual’s current self-assessed state of health and then introduced them to the particular illness descriptions that constituted the focus of subsequent questions. Part B introduced the general idea of WTP for health benefits before eliciting a series of WTP values for different scenarios. These scenarios were varied between four different subsamples in order to provide broader coverage and allow certain tests of conformity (or otherwise) with standard assumptions. Respondents were allocated at random to one of the four subsamples, and the fact that Part A and the first question in Part B were common to all respondents allowed us to check for the comparability of those different subsamples. Part C consisted of four questions eliciting various risk trade-offs designed both to test the standard assumptions invoked for this part of the QALY computation and to provide a QoL/health-state index measure for the illness state under consideration. The final part of the questionnaire consisted of a series of questions collecting socioeconomic/demographic information. This allowed further checks of comparability between subsamples.

Implementation and results

There were four versions of the questionnaire, one for each of the four subsamples. A copy of Version 2, together with the relevant supporting material, is included in Appendix 6. The ways in which the other versions differed from Version 2 will be explained in the text that follows.

But first, the sample. In order to make the most of the project budget, sampling for this part of the project involved revisiting many of those who had previously participated in the ‘relativities’ part of the study and then enlarging the sample opportunistically from that base. Thus, this was a ‘convenience’ sample; and while it covered the full range of educational level, income and social class, it was *not* intended to be representative of the population as a whole. What is important for present purposes is that within this sample, respondents were allocated at random between the four versions of the questionnaire: our objective was to examine within-subject consistency and make between-subsample comparisons. If it is judged that the survey instruments pass the tests conducted on this basis, it will require a larger and more representative sample to provide estimates suitable for public policy.

The four versions were as follows. Versions 1 and 2 focused on scenarios about stomach illness while Versions 3 and 4 were centred on scenarios about head pain. As indicated earlier, within each version there were three durations of illness: 3 months followed by return to respondent's current state of health; 12 months followed by return to respondent's current state of health; and a chronic condition where the illness lasted for the rest of the respondent's life. Each scenario was described on a separate card: the six cards are reproduced as part of the supporting materials in Appendix 7.

The key way in which Version 1 differed from Version 2, and likewise the way in which Version 3 differed from Version 4, related to the questions involving the risk of an illness: Versions 1 and 3 asked about WTP to eliminate 10% risks of the three durations of illness, while Versions 2 and 4 asked about eliminating 5% risks.

The total number of questionnaires was 409. However, after a check, six of these were excluded on the grounds that they contained very few answers to the main questions and offered little or no possibility of contributing usefully to the analysis.

The distribution of the 403 between the four versions is shown *Table 26*.

TABLE 26 *The distribution of versions*

Version	Illness	%	Total
1	Stomach pain; chance 10%	26	105
2	Stomach pain; chance 5%	27	108
3	Head pain; chance 10%	25	99
4	Head pain; chance 5%	23	91
Total			403

TABLE 27 *The distribution of answers for current state of health*

	No problems	Occasional minor problems	Frequent but minor problems	Quite a lot of difficulties	Very severe difficulties
Mobility	309 (77%)	45 (11%)	23 (6%)	21 (5%)	5 (1%)
Self-care	376 (93%)	14 (3%)	7 (2%)	4 (1%)	2 (0.5%)
Usual activities	317 (79%)	44 (11%)	20 (5%)	18 (4%)	4 (1%)
Pain and discomfort	204 (51%)	125 (31%)	35 (9%)	32 (8%)	7 (2%)
Anxiety and depression	300 (75%)	76 (19%)	16 (4%)	8 (2%)	2 (1%)

We now turn to the main results from the key parts of the questionnaire, starting with Part A.

Part A: current health status and ranking of illness descriptions

At the beginning of the questionnaire, respondents were asked to complete a modified (actually, expanded) form of EuroQol 5 dimensions (EQ-5D) (five levels for each dimension rather than the usual three levels) – we shall refer to our version as EQ-5D+. Then they were asked to rate their own health on a visual analogue 'thermometer' scale, where 100 was labelled 'as good as it could be for someone of your age' and 0 was 'as bad as being dead'.

These questions were included more for purposes of 'warming up' respondents – getting them to think about what we mean by health, and giving them some tasks which involved them from the outset. They also allow us to check that there were no obvious health disparities between the respondents across the four versions – and indeed there were no significant differences between the distributions of responses to these questions across the subsamples. (Additional checks on the distributions of age and gender also showed no significant differences between the four subsamples in those respects.)

Table 27 gives an overall summary of the responses to the EQ-5D+ questions.

To check whether there was broad correspondence between EQ-5D+ responses and the VAS, we grouped EQ-5D+ responses into four categories by adding up the rank on each dimension. So someone who reported themselves as problem free, i.e. as 11111, got a sum score of 5. Someone who reported just one 2 got a score of 6. Someone who

reported themselves as 21232 was scored as 10; and so on. *Table 28* gives the main VAS statistics for each of four EQ-5D+ categories.

Comparisons across the groups using parametric and non-parametric tests [analysis of variance (ANOVA) and Kruskal–Wallis] reject the null hypothesis in favour of the alternative that there are significant differences: that is, VAS scores are correlated with EQ-5D+ responses and VAS scores are as we should expect. In addition, both appear to correlate with age – the older people are, the worse their self-reported health.

The questionnaire next asked each respondent to consider the various illness descriptions reproduced in Appendix 7 – 3 months of stomach illness, 12 months of stomach illness, the stomach illness for the rest of life, 3 months of head pain episodes, 12 months of head pain episodes, the head pain episodes for the rest of life, and sudden painless death – and to rank these seven scenarios from least bad down to worst. The main point of this exercise was to introduce people to the descriptions and to give them a task intended to get them to read them fairly carefully.

Piloting had shown that it was quite demanding simply to give these seven cards to respondents and ask them to rank them. So the actual procedure involved giving them four cards to start with – the 3-month and 12-month durations of both the stomach and the head pain symptoms – and ask them first to rank those. Once these had been ranked, each respondent was given the two ‘lifetime’ durations and was asked to add those to the ranking. Finally, they were asked to locate the ‘sudden painless death’ card by putting it above any of the other descriptions the respondent considered worse than death, or by putting it at the bottom of the ranking if the respondent considered death to be worse than all of the six illness scenarios.

There is no single obviously ‘right’ ranking over all seven cards, but at the very least we might expect that within ‘stomach’ and within ‘head’ we should find the lifetime duration to be ranked worst and the 3-month duration to be ranked least bad. However, 21% of respondents failed to satisfy that expectation for at least one body area. This is arguably rather disappointing.

Subsequent analysis shows that the ‘aberrant’ 21% did not give systematically different answers to the sets of questions that followed. Our conjecture is that although there was evidence of some confusion when asked to process these descriptions on first presentation, the later questions, which focused on subsets of no more than three at a time, did not pose the same difficulties. Still, one clear message is that if a large-scale representative survey were to be undertaken in the future, and if this were to involve a significant ranking task (especially near the beginning), extra time and care should be given to allow respondents to process the task. However, in the present feasibility study, the ranking exercise was intended mainly as a warm-up/familiarisation task, and respondents’ particular rankings play no role in the subsequent analysis.

Part B: willingness-to-pay questions

Having completed the illness ranking exercise, respondents were introduced to the idea of WTP to prevent/reduce illness with a ‘practice’ question, common to all versions. This question asked about WTP to reduce the duration of a sore throat while on holiday from 3 days to 1 day. The overall mean WTP was just over £40, with a median of £20. There were no differences between the distributions of responses across the four subsamples – again, some reassurance that the assignment of respondents to version was such as to allow comparability across subsamples.

TABLE 28 Main statistics for VAS values between EQ-5D groups

	Sum of ranks=5 (IIIII)	Sum of ranks=6	Sum of ranks=7/8/9	Sum of ranks=10+
<i>n</i>	169	87	86	61
Mean	90.3	83.9	79.67	63.42
Median	94	85	80	70
Standard deviation	12.52	13.10	12.83	16.56

Respondents then embarked on a series of five WTP questions. These are labelled B4, B5, B6, B7 and B8 in the questionnaire, and this report will stick with those labels. B4 asked respondents what they would pay for 'a simple, safe and painless cure that would avoid' the (otherwise) certainty of the 3-month illness (stomach in Versions 1 and 2, head in Versions 3 and 4); B5 asked for WTP to avoid the certainty of the 12-month illness; B6 asked for WTP to eliminate either a 10% (Versions 1 and 3) or a 5% (Versions 2 and 4) risk of the 3-month illness; B7 asked the same for the 12-month illness; and B8 asked for WTP to eliminate the 10%/5% risk of the respondent suffering the illness for the rest of their life.

The precise wording of these questions can be seen in the copy of Version 2 in Appendix 6. The essential features were as follows. First, respondents were asked to suppose that their income would be unaffected and to focus just on how the illness would affect their health, with health being explicitly framed with reference to the EQ-5D+ dimensions they had considered in the first question in Part A.

In each of both B4 and B5, respondents were first asked whether avoiding the health effects of the illness would be worth at least something, even if it were only a few pence. If they responded that it would not even be worth paying a few pence, they were asked to say why they felt that way, and their response was noted by the interviewer. However – and especially for B4 and B5 (the questions expressed in terms of certainty) – the overwhelming majority said that they *were* willing to pay something. These respondents were then given a small pack of cards, each of which had a different sum of money printed on it. (The amounts were: £1, £5, £25, £50, £100, £250, £500, £1000, £2000, £3000, £5000, £10,000, £20,000, £50,000, £100,000 and £1M.) Once the pack had been shuffled, they were asked to sort the cards into (up to) three piles: those amounts they definitely *would* pay for an instant cure that would avoid all aspects of the illness; those amounts they definitely would *not* be prepared to pay; and any amounts about which they were unsure (initially, at least) whether or not they would pay.

After initially sorting all of the money amounts, they were asked to think about any amounts they had placed in the 'unsure' pile and were invited to relocate them (if they wished – although there was no pressure to do so) into either of the other

two piles. The interviewer then recorded the highest amount in the 'definitely would pay' pile and the lowest amount in the 'definitely would not pay' pile, reminded the respondent of the highest amount in the 'definitely would pay' pile and offered them the opportunity to nominate any higher amount they felt they would be prepared to pay. If the respondent did nominate a higher amount, that amount was recorded and was used in the subsequent analysis; if the respondent did not nominate a higher amount, the largest amount in the 'definitely would pay' pile was used in the analysis.

Exactly the same procedure – and as far as possible, the same wording – was used for B6, B7 and B8: the only substantive difference was that respondents were now being asked to suppose that a test had shown there was a 5% (alternatively, 10%) chance of developing the illness in question and were asked for their WTP to eliminate that risk.

Avoiding/preventing the certainty of illness

It should be noted that for questions B4 (3-month certainty) and B5 (12-month certainty), there were very few people giving zero responses. For the 12-month stomach (S) condition in B5, there were only two, while only seven others gave a WTP below £25; for head pain (H), the figures were one zero response and four others below £25. This contrasts with many WTP surveys where there is often a significant proportion of 'protest' zero responses given by people who feel that the good in question ought to be provided without specific charge and/or that money values are inappropriate. In this sample there appeared to be no such reaction to the avoidance of an otherwise certain illness: on the contrary, there was general recognition that the benefits offered in B4 and B5 were significant, and there was a widespread preparedness to indicate that with positive WTP responses.

Besides establishing whether there was or was not some general willingness to state a monetary value to avoid certain illnesses, questions B4 and B5 were included for two other reasons.

The first objective was to see whether spending some period(s) in one health state was regarded as significantly worse than spending the same period(s) in the other health state, as reflected by some significant difference between WTP to avoid a given duration in each state. (The intention was

to compare this judgement with the health-state indices elicited via later questions.)

The second objective was to examine the relationship between WTP and the duration of the illness.

For this latter purpose, the head pain cases are the most clear-cut. As can be seen from the illness description cards V and A in Appendix 7, the *only* difference between the two is the duration: V lasts for 3 months and A lasts for 12 months. But how might this difference be reflected in WTP amounts?

There are various possibilities, not all working in the same direction. First, even though the periods of time are relatively short and reasonably close to the present, it might be that the future is discounted, so that the second, third and fourth lots of 3 months that constitute the last 9 months of the 12-month period are each given progressively lower weights, with the result that experiencing the health state for 12 months is regarded as less than four times as bad as experiencing it for 3 months. Moreover, there are 'adaptation' arguments that suggest that people may get used to being ill and adapt their behaviour accordingly, so that the loss of welfare after adaptation is less than it is initially – with the result that the welfare loss over 12 months is less than four times the loss over 3 months.

In addition, and working in the same direction, it may be that paying four times as much represents more than four times the 'sacrifice' as financial budget limitations bite increasingly hard – with the result that WTP to avoid the 12-month illness is substantially less than WTP to avoid the 3-month illness.

Against that, there is the possibility that a longer period of illness becomes increasingly difficult to accommodate and/or tolerate. Two somewhat different things might be at work here. First, there may simply be a sense of being ground down by continuing ill health. Second, while it may be possible to rearrange one's life and work to some extent to cope with shorter periods of illness – putting off a holiday, for example, or postponing some tasks for a while – it may be increasingly difficult to work around a longer period of illness.

It was not obvious a priori which tendencies would be predominant, and it was not intended that this study design would be able to unscramble all of the various forces and factors. The more limited

aims were these: first, to see whether the pattern of money values corresponds with the simplest QALY model, where QALY losses are supposed to be strictly proportional to the time spent in an illness state, or whether there is some other pattern; and second, whether the pattern of money values corresponds with the trade-offs made in Part C.

The same aims apply to the stomach illness scenarios. These were slightly more complex: although the recurrent bouts of stomach discomfort and sickness were of the same intensity and frequency for both the 3-month and the 12-month scenarios (and in that respect the latter might be regarded as four times the former), they began with illnesses of different severities, i.e. 3–4 days of stomach pains, diarrhoea and vomiting at the start of the 3-month illness compared with 7 days of *severe* stomach pains, diarrhoea, vomiting and fever at the beginning of the 12-month illness. However, even if the latter cannot be known to represent exactly four times as much loss of well-being as the former, we should still expect to see clear differentiation between the two, and the money values and Part C trade-offs can still be compared.

We shall come to those Part C trade-offs later. But first, we consider the issue of whether the stomach condition (S) generated significantly different WTP responses than the head pain condition (H). *Table 29* gives the summary statistics.

The presence of outliers increases standard errors and makes it more difficult to reject via *t*-tests the nulls of equal means between S and H, although these nulls *are* just rejected at the 5% level (2-tailed) for both 3 months and 12 months. The non-parametric (Mann–Whitney) tests reject the null (easily) at the 1% level in both cases. On this evidence, members of our sample seemed to regard the H state as worse than the S state.

What about sensitivity to duration? The 12-month to 3-month ratio of means is 2.305:1 for S and 2.174:1 for H – both well short of the 4:1 ratio of duration. The ratios of medians are closer, i.e. 3.33:1 for S and 4:1 for H. So there is *some* sensitivity, but it is as if there is either some discounting of the longer duration or some effect of budget constraints (or some combination of both).

If we take the *difference* between the B5 and B4 responses for each individual, we find a similar overall pattern for both S and H. For S, 11 go the wrong way (i.e. WTP strictly more to prevent

TABLE 29 Summary statistics for B4 and B5

		3 months		12 months	
		Stomach	Head	Stomach	Head
<i>n</i>	Valid	212	189	213	190
	Missing	1	1	0	0
Mean WTP (£)		810.27	1495.88	1867.37	3252.35
Median WTP (£)		150.00	250.00	500.00	1000.00
Standard deviation		1856.70	4658.29	4769.52	8254.29

3 months than 12 months), 45 give exactly the same answer to both questions (i.e. difference = 0), and 156 give a higher WTP to prevent 12 months. For H, the corresponding breakdown is 10:32:147. Were we to focus exclusively on the 303 people who gave a strictly higher response to B5 than to B4, the ratio of means for S would be 2235.83:687.21 which is 3.25:1, and for H the ratio would be 3870.44:1086.78, i.e. 3.56:1. Thus, even taking the subset of responses which satisfy the (minimum) requirement that the respondent places a strictly higher value on the avoidance of the 12-month illness than on the avoidance of the 3-month illness, the means still suggest a less-than-proportional relationship between WTP and duration.

With such relatively short durations as 3 and 12 months, it might seem implausible that there is substantial discounting on time preference grounds alone – in which case, the prime suspect might be the impact of budget constraints. If this were the main cause, we might expect the effect to be reduced by moving away from certainty and asking questions about reducing risks of adverse outcomes. This brings us to questions B6, B7 and B8.

Eliminating the risk of illness

Tables 30–32 report the key statistics for the questions which asked respondents for their WTP to eliminate either a 10% risk or a 5% risk of suffering the specified illness (B6 = 3-month risk, B7 = 12-month risk and B8 = lifelong risk).

The first issue is whether responses were sensitive to the size of risk reduction. If such questions are to provide a sound basis for public policy, we should ideally like to see the values for eliminating the 10% risks being close to double the values for eliminating the corresponding 5% risks; and at the very least, we should expect the former to be significantly higher than the latter.

However, Tables 30–32 show that there was very limited sensitivity to the size of the risk being eliminated. Out of the six 10% versus 5% between-sample comparisons – three for S and three for H – none of the *t*-tests showed a statistically significant difference between means. Indeed, and rather worryingly, the mean values for eliminating the 5% risk of H were in fact all *higher* than the corresponding means for eliminating the 10% risk. To some extent, this can be explained in terms of a couple of high outliers in the 5% H subsample (note the much larger standard deviations). However, there are other worrying signs – not least that the medians for the H subsamples, in particular, show insufficient sensitivity (and in B8, actually go the wrong way); and out of the six relevant Mann–Whitney tests, only one, for the 12-month stomach illness, showed a difference that was statistically significant at the 5% level.

We can also compare on a within-subject basis the difference between what a respondent said they were willing to pay to prevent the certainty of an illness and what they said they were willing to pay to eliminate a 10% or 5% risk of the same illness. Generally, about 80% of the sample were willing to pay strictly more to prevent the certainty than to eliminate the risk, which would seem to be a minimum requirement. However, that still leaves about 12% who gave the same WTP, and about 8% who said they were willing to pay *more* to eliminate the risk than to prevent the certainty.

What about within-subject sensitivity to duration, given the baseline risk? For each of the four versions, we can examine at the individual level the difference between the B8 (lifelong) response and the B7 (12-month) response, and also between the B7 and B6 (3-month) responses.

Running *t*-tests of whether these differences are significantly greater than zero gives a clear picture for the B8 versus B7 differences – all reject the

TABLE 30 B6 – eliminating risk of the 3-month illness

	Stomach		Head	
	10%	5%	10%	5%
<i>n</i>	104	106	98	90
Mean WTP (£)	375.83	230.44	403.36	477.34
Median WTP (£)	25.00	25.00	50.00	50.00
Standard deviation	1133.79	1030.80	1283.01	2221.59

TABLE 31 B7 – eliminating risk of the 12-month illness

	Stomach		Head	
	10%	5%	10%	5%
<i>n</i>	105	106	99	90
Mean WTP (£)	914.44	451.70	623.87	877.29
Median WTP (£)	100.00	50.00	150.00	100.00
Standard deviation	3364.82	1666.58	1460.66	2337.15

TABLE 32 B8 – eliminating risk of the lifelong illness

	Stomach		Head		
	10%	5%	10%	5%	5% ^a
<i>n</i>	104	108	98	86	86
Mean WTP (£)	3235.39	1883.42	3008.94	14,249.64	3203.13
Median WTP (£)	325.00	250.00	600.00	750.00	750.00
Standard deviation	9565.18	4145.36	6471.16	107,630.34	6637.52

a Replaces one observation of £1M with £50,000.

null at the 1% level. However, the means and medians for B8 are mostly only three to five times the values for B7, whereas average remaining life expectancy is much more than 3–5 years: the mean and median ages of the sample were 51.37 and 52 respectively, so that average remaining life expectancy would be in the region of 30 years.

The differences are predominantly in the right direction for the B7 versus B6 comparison, but here there are many more zero differences – 160 across the sample as a whole, many but not all the result of zero WTP to eliminate these risks – so the null is not rejected for Version 4 and is on the borderline for Version 1. The ratios of the means are 2.43:1, 1.96:1, 1.55:1 and 1.84:1 for the four versions – all well short of the 4:1 ratio of duration, and overall lower than the corresponding ratios for B4 and B5, despite the fact that budget

considerations might have been expected to have exerted greater constraints on the B5:B4 ratios. The ratios of medians are somewhat higher, at 4:1, 2:1, 3:1 and 2:1, but mostly still fall short of the ratio of duration and are also more muted than in the B5:B4 comparisons. This may suggest that it is not simply an issue of budget constraints but that the use of probabilities introduces further ‘noise’ and diminished sensitivity into people’s patterns of response.

A possible side effect of the reduced sensitivity to questions involving these risks was that the difference between S and H appeared to be somewhat attenuated. There are three comparisons for the 10% baseline, and another three for the 5% baseline: none of the six *t*-tests showed a statistically significant difference, and only one of

the Mann–Whitney tests came close ($p = 0.056$ in the case of the 5% risk of the 3-month illness).

Thus, despite the theoretical advantages of using risk reduction questions to elicit values, the evidence from this feasibility study seems to suggest that those questions do not, in practice, show the kind of sensitivity, either to the size of the risk reduction or to the duration of the illness, that would be desirable as a basis for a robust value of a QALY.

If probabilities seem to pose additional problems for WTP questions, what are the implications for questions that use probabilities to express trade-offs between health states and durations? The study explored this issue in Part C of the questionnaire, to which we now turn.

Part C: standard gamble questions

There were four questions in this part of the questionnaire. The basic structure was the same for all four. In each case, respondents were asked to consider two alternatives. On the left of a showcard (see the supplementary material in Appendix 8) was displayed some prospect that would be faced with certainty – for example, in the first of these questions, that prospect was the certainty of 3 months of either S or H. On the right of the showcard was an uncertain prospect with two possible outcomes: the good outcome, which was in all cases the respondent's current health, with no ill effects; and the bad outcome, which varied from one question to another, but which was always worse than the certain prospect on the left – for example, in the first of these questions it was 12 months of either S or H.

So the uncertain alternative offered some chance of being in a better state than the certain prospect, but also entailed some chance of being worse off. What each respondent was asked to identify were the probabilities of being better or worse off that would make them feel that there was nothing to choose between the two alternatives. In all four questions, these probabilities were elicited by an iterative process. The full details can be found in the body of the questionnaire in Appendix 6, but the essential idea was as follows.

Initially, the respondent was asked to choose between the certain prospect (of 3 months in S, say) and a treatment whose outcome was uncertain

but which offered a 90% chance of success (which meant continuing in current health, with no ill effects) and a 10% chance of failure (resulting in S lasting for 12 months rather than 3 months). The next stage of the question depended on the answer to that first choice. If they had preferred the certain prospect to the 90%:10% alternative, the chances were changed to make the uncertain alternative more attractive – this time offering a 99% chance of success alongside a 1% chance of failure – and the respondent was asked to choose afresh. On the other hand, if they had preferred the uncertain alternative, the chances were changed to 50%:50% to make that prospect less attractive, and the respondent was then asked to choose on this basis.

By altering the chances on the right hand side in response to each choice, it was possible to home in on a pair of probabilities which balanced the two alternatives in terms of the respondent's preferences.

As indicated above, the first question of this kind – labelled C and administered in conjunction with the relevant showcard – involved balancing the certainty of a 3-month illness against a risky treatment offering instant recovery to current health if successful but suffering for 12 months if this treatment failed.

The next question – labelled D in the questionnaire – involved the certainty of 12 months' illness versus a risky treatment offering instant recovery to current health if successful but suffering for the rest of their lifetime if this treatment failed.

Question E juxtaposed the certainty of the lifelong condition versus the prospect of current health if the uncertain treatment succeeded but (painless immediate) death if it failed.

Finally, Question F was the 'complement' of D. That is, the certain prospect involved 1 year in current health, which would then be followed for sure by the rest of their life in either S or H; while the risky prospect offered current health if it succeeded, but if it failed it entailed S or H starting immediately (i.e. this meant losing the first year in current health).

Question C was intended to explore the relationship between the subjective loss entailed by the 3-month illness compared with the loss of well-being associated with suffering the condition for four times as long. As no money is involved

in these types of question, the focus is upon how much worse the 12-month duration is compared with the 3-month period of illness. If (as a simple QALY calculation supposes) the 12-month illness involves four times as much QALY loss as the 3-month illness, a respondent behaving broadly according to the QALY model should feel that the two alternatives are evenly balanced when the risk of failure is 0.25: that is, the 0.75 chance of avoiding the 3 months of illness would be regarded as exactly balancing a 0.25 chance of suffering an extra 9 months.

As Table 33 shows, a relatively small proportion (overall, about 15%) of each subsample gave that answer. The majority (about 60% overall) were willing to accept only a smaller risk. Neither parametric nor non-parametric tests suggest any difference between the distributions of responses for S and H. On average, the 12-month illness appears to be regarded as five or six times as bad as the 3-month duration; while the interpretation based on the medians is that most people regard the loss of well-being (or to use decision theoretic terminology, utility) from the 12-month illness as being between seven and 14 times the loss involved in a 3-month illness.

If we took these ratios as a reliable reflection of individual attitudes to duration, it would suggest that considerations of the difficulty of adapting to anything more than a fairly short illness and/or the 'intolerability' of continuing ill health outweigh any 'time discounting'. These results contrast with the relativities suggested by the responses to B4 and B5, where the ratios of means were only slightly above 2:1 and the ratios of medians were between 3:1 and 4:1. So if the SG responses do reflect preferences, the implication is that budget constraint effects are really quite strong. However, an alternative possibility is that framing these questions in terms of 'chances' and requiring

TABLE 33 The statistics for Standard gamble C

	Stomach	Head
<i>n</i>	202	179
Failure risk <0.25	130	102
Failure risk =0.25	29	28
Failure risk >0.25	43	49
Mean	0.177	0.207
Median	0.075	0.15
Standard deviation	0.19	0.21

respondents to think probabilistically and confront risk and uncertainty may prompt excessively cautious responses, exerting a downward influence on the chances of failure they would accept, and seeming to inflate the magnitude of the loss entailed by 12 months as compared with 3 months. As will be seen shortly, there is *some* evidence from this study which is consistent with this latter interpretation; but the evidence is decidedly mixed, as is shown by the next question.

Question D involved risking the illness for the rest of their life in order to avoid the certainty of 12 months of the illness. As the lifelong illness involves more years of health loss for those with the greatest remaining life expectancy, one might suppose that the acceptable risk of failure should be lower for younger people. Table 34 divides the full sample into three roughly equal age groups and then reports the mean and median risks of failure for each age group for both S and H.

However, for the most part there are no statistically significant differences between age groups within S or within H, whether judged by *t*-test or by Mann–Whitney test. The only exception is the comparison between the middle and older age groups for S: but there, the difference is in the *opposite* direction to the expected one – that is, older respondents are willing to take less risk, even though they would suffer for fewer years.

The lack of – or even, perverse – sensitivity of response to age is a discouraging result. But here, by contrast with Question C, the means appear to *underweight* the magnitude of loss: given that average remaining life expectancy is in the region

TABLE 34 The statistics for Standard gamble D

	Mean risk of failure	Median risk of failure
S		
Age ≤ 43 (<i>n</i> = 61)	0.101	0.055
Age > 43 but < 59 (<i>n</i> = 66)	0.137	0.060
Age ≥ 59 (<i>n</i> = 73)	0.084	0.025
All S	0.108	0.055
H		
Age ≤ 43 (<i>n</i> = 62)	0.117	0.045
Age > 43 but < 59 (<i>n</i> = 56)	0.123	0.030
Age ≥ 59 (<i>n</i> = 60)	0.150	0.050
All H	0.129	0.035

of 30 years, overall means of 0.108 (for S) and 0.129 (for H) imply that suffering the conditions for the rest of life would only be between 8 and 10 times as bad as suffering for 1 year – a result that seems implausible unless we assume respondents to be discounting the future very heavily and/or to be anticipating considerable adaptation. Another possibility is that the arithmetic mean is not a particularly good measure of central tendency in these cases: the medians of 0.055 and 0.035 – implying that the lifetime illness is between 18 and 28 times as bad as the 12-month illness – may constitute a better reflection of the ‘typical’ respondent.

Before considering Question E, it may be useful to look at the results from Question F. As mentioned above, Question F is the ‘complement’ of Question D. For both questions, the possible outcomes of the risky treatment are either to continue in current health or else to suffer lifelong illness from now on. In Question D, the certainty is 12-months’ illness followed by a return to current health for the rest of life, so that the possible gain from the risky treatment is to avoid the illness for the coming year, while the possible loss is to drop from current health for the rest of life *after* this year to illness for that remaining lifetime.

In Question F, the certainty is that the respondent will remain in current health this year, but the onset of illness cannot be delayed beyond that and so the respondent will then be ill for the rest of their life. Thus, the potential gain from the risky treatment would be to avoid illness for all life after next year and instead spend those years in the health currently expected (i.e. the mirror image of the potential loss in Question D), while the possible loss in Question F is that instead of spending the next year in current health, the effects of the illness start now (i.e. the mirror image of the potential gain in Question D). On standard assumptions, whatever chance of failure makes the respondent feel that the alternatives are finely balanced in

Question D should be the chance of success that would make things finely balanced in Question F. Put another way, for each respondent the sum of the two risks of failure should come to 1. *Table 35* reports the vital statistics for Question F and for the D + F sums.

The null hypothesis that $D + F = 1$ is rejected at the 0.1% level for both S and H. To give some rough idea of the distributions, consider *Table 36*, where the observations are grouped in four ranges.

It is not clear exactly what this result signifies. If a similar exercise has been conducted in previous work, we are not aware of it; and so interpretation must be cautious and speculative. However, one possible interpretation is that many people are averse to the possible bad outcome of a risky prospect but do not know exactly how averse they are, and therefore in one or both questions they respond with extra caution, thus producing at least one and possibly two understatements of the downside risk they would accept.

While this might fit with the tentative suggestion that respondents exhibited excessive caution in Question C, it does not fit so neatly with the responses in Question D, which might be seen as reflecting a propensity by a significant proportion of the sample to take too much risk when the downside involves suffering the illness for the rest of their life.

On the other hand, it does seem likely that the risks people were willing to take in Question F were very conservative. An ‘average’ respondent with 30 years’ remaining life expectancy is facing, on the left-hand side, the certainty of spending the last 29 of those years in either S or H; but the uncertain treatment on the right-hand side offers some chance of avoiding that and instead spending all of the rest of their life in the health they currently expect – the downside risk being that if the uncertain treatment fails they lose

TABLE 35 The statistics for Standard gamble F and for the sums of D and F

	Stomach		Head	
	F	D+F	F	D+F
<i>n</i>	202	195	180	172
Mean	0.474	0.583	0.520	0.653
Median	0.550	0.575	0.550	0.676
Standard deviation	0.33	0.39	0.33	0.41

TABLE 36 Indicative distributions for D + F

	Stomach	Head
$D + F \leq 0.5$	87	63
$0.5 < D + F < 0.9$	60	51
$0.9 \leq D + F \leq 1.10$	35	41
$D + F > 1.10$	13	17

the coming year of current health. The median response – wanting a 45% chance of avoiding the chronic illness in order to be prepared to take a 55% chance of suffering the illness immediately – appears at first pass either to place a very high value on the next year relative to the rest of life (not reflected in their Question D responses) or to reflect misunderstanding/confusion. In any event, it may prompt us to treat with caution other data emerging from SG questions.

This brings us to Question E, which is a vital component in any attempt to infer a monetary value of a QALY from the present survey.

The idea of Question E was to get each individual's assessment of the health state (S or H) in the form of a 'health-state index' number. The underlying assumption here is that current health is indexed at 1 and death at 0 and that an individual's index number for either S or H is independent of the number of years spent in that state relative to spending those years in current health or else being dead for that time. So the loss of health (relative to 1) is assumed to be revealed by the risk of death that would make the individual indifferent between the certainty of S/H for the rest of their life and the uncertain cure-or-kill treatment. The results for this are portrayed in *Table 37*.

Although the mean for H was higher than for S, the medians were the same, and neither a *t*-test nor a Mann–Whitney test registered any significant difference. The picture here, then, is more similar to the one emerging from the risk reduction WTP questions B6–B8 than from the B4 and B5

TABLE 37 The statistics for Standard gamble E

	Stomach	Head
<i>n</i>	196	177
Mean	0.104	0.144
Median	0.025	0.025
Standard deviation	0.17	0.25

questions asking about WTP to avoid the certainty of the illnesses, which suggested that H was regarded as significantly worse than S.

Using the willingness-to-pay and standard gamble data to derive a monetary value of a QALY

In order to get an estimate of each individual's monetary value of a QALY, our procedure involves converting an illness description into a QALY loss by using the individual's response to Question E, and then combining that with their money value for avoiding that loss, taken from the WTP questions.

There are various ways we could combine individuals' responses to different questions, but perhaps the simplest to explain is to take the 12-month illness, compute the QALY loss involved in that, and combine that with WTP to avoid that loss as stated in response to B5. It will be seen that the issues raised by this route are so fundamental that it is really rather superfluous to pursue the other variants, none of which circumvent those fundamental problems. So we focus on this route, which works as follows.

Consider someone who is just willing to accept a 5% risk of death in Question E. This is taken to signify that the illness state is indexed at 0.95 on the scale that runs from 1 (normal health) to 0 (dead, or as bad as being dead). In other words, each year spent in that state is rated at 0.95 of a normally healthy year and thus represents a loss of 0.05 (i.e. one-twentieth) of a QALY. If that same individual states in Question B5 that they are willing to pay £800 to avoid that loss, then 20 people like this respondent would, collectively, be prepared to pay £16,000 to avoid health losses that add up to one QALY (i.e. 20×0.05). Thus, this individual's responses amount to saying that for a population consisting of similar people, the money value of a QALY is judged to be £16,000. By calculating the 'value of a QALY' implied by each individual's responses combined in this way, we can derive mean and median values for this sample.

The potential problem, of course, is that if someone says they are only willing to take a 1 in 100,000 risk (or less) of death in Question E, their WTP to avoid the 12-month illness is multiplied

by 100,000, potentially giving an astronomical figure for the value of a QALY. (Someone who says they are only willing to take a 1 in 100,000 risk of death to avoid the chronic illness state is taken to be indexing that health state at 0.99999, i.e. a year spent in that state is taken to amount to the loss of 0.00001 of a QALY. So if 100,000 such people were each willing to pay, say, £300 to avoid the 12-month illness, they would between them be paying £30M and their combined benefit would add up to just one QALY.) Sure enough, we find that 115 respondents in total – 59 in the S subsample and 56 in the H subsample – give responses to Questions E and to B5 which, in combination, imply values of a QALY of more than £1M. With some of these combinations generating values of thousands of millions of pounds, the mean values for a QALY are £3 × 10⁸ from the S subsample and £7 × 10⁸ from the H subsample. The medians are somewhat more terrestrial: £26,666.67 and £57,142.86 respectively. On the grounds that there are no statistically significant differences between the two subsample distributions, we might pool them to obtain an overall mean of £5 × 10⁸ and an overall median of £40,000.

It has been widely accepted that when such values are being used to guide public policy, it is the mean figure which should be used as the best indicator of social welfare. However, clearly, a mean value for a QALY of £5 × 10⁸, or even a figure one-thousandth as big as that (i.e. £500,000), would be totally anomalous in a world where the VPF is about £1.5M and where this, in the context of road accident fatalities, represents preventing a death which on average entails the loss of about 40 years of life expectancy.

Nevertheless, applying the procedure in the way described does generate the kinds of extremely skewed distributions reported, and that does produce means which are pulled up by some extremely high individual figures.

This raises the question of whether there is something that could be done – some other way of managing the procedure or analysing the data, perhaps – that might give figures more compatible with those in use in other areas of public health and safety policy?

As a start to answering this question, recall what is involved in a study to elicit the value of preventing a (statistical) fatality or the value of preventing a particular (statistical) injury. Typically, respondents

are presented with a description of a specified physical outcome – death, or a particular injury description – and are asked for their WTP to reduce their risk of experiencing this outcome by some given amount. We then effectively sum the responses over a representative population big enough that their individual risk reductions add up to preventing the (statistical) expectation of one death or one injury of the particular type.

In the course of this procedure, no attempt is made at the individual level to ascertain the number of QALYs each individual perceives they would lose in the event of being killed or injured. In principle, this *could* be done. Were things to be done that way, it might very well be that the combination of some individuals giving high WTP with responses by them implying very small individual QALY scores would produce a similar highly-skewed distribution and upward pull on means. But as things stand, it has not been the usual practice. So, in the light of the way in which things actually have been done, what might be the implication for the present study?

Take the case of a non-fatal road injury, which is a closer analogue than a fatality to the illness descriptions used in the present study. Call this road injury 'J'. Typically, respondents have been given descriptions of what J involves, have been asked to think how such an injury would impact upon their life, and have then either been asked for their WTP to prevent/reduce the risk of such an injury or expressed some relativity between the injury and death. In this latter case, the value of preventing each J is then 'pegged' against the VPF according to the average relativity expressed between J and death. But however it is arrived at, for the sake of example, let us suppose that the process generates an average money value for preventing injury J of £60,000.

If one were to want to go further and infer a value of a QALY from the established value of preventing injury J, the most obvious way of doing so would be to undertake a survey eliciting from a cross-section of the population their judgement of the QALY loss they would suffer if they sustained injury J. Again, for the sake of example, suppose that responses to such a survey ranged quite widely from one individual to another but that the average judged QALY loss turned out to be 2.5 QALYs.

On this basis, a public body undertaking a road safety scheme costing £600,000 which is expected

to prevent 10 cases of injury J could (just) justify that expenditure on cost-benefit grounds (assuming all other projects yielding an excess of benefit over cost were also undertaken). And although this body could not say in advance which particular individuals would benefit from the scheme by avoiding injuries they would otherwise have suffered, it could assert that on average the benefit would amount to preventing a loss of 25 QALYs, translating to an average value of £24,000 per QALY.

We can mimic that kind of calculation on the basis of the data collected in the feasibility study if it is processed in the following way.

First, we have reported in *Table 29* the WTP-based mean values for avoiding the 12-month durations of S and H: respectively, £1867.37 and £3252.35. Let us round these figures to £1870 and £3250. Had our sample been a large, representative cross-section of the population, it could have been argued that health-care resources allocated to preventing/curing such illnesses could (just) be justified if they cost those amounts for every 12-month episode prevented/cured.

Then we might ask what QALY gains would result from such expenditure. We cannot know which particular individuals will benefit, but on the basis of our sample members' responses to Question E as summarised in *Table 37*, together with the standard QALY procedure for combining health-state indices with periods of time spent in a health state, we could say that, on average, each avoided case of 12-month S would avoid a loss of 0.104 of a QALY, while each avoided case of 12-month H would generate an average benefit of 0.144 of a QALY.

Putting the relevant figures together, we would have an expenditure of £1870 on 12-month S, bringing an average QALY benefit of 0.104, which translates to a money value of about £17,980 per QALY, while each £3250 spent preventing 12-month H would 'save' 0.144 QALYs, implying a value per QALY of approximately £22,570.

Processed this way, the data obtained via two different health-state descriptions produce figures reasonably close to each other on either side of £20,000 per QALY. But why are these figures so very different from the astronomical means based on the value-per-QALY figures generated by combining Question E and B5 responses at the level of each individual? And which – if either – is the correct basis for deriving an estimate?

The answer to the first of these questions is fairly straightforward and can be illustrated by an example (which may also be helpful when trying to answer the second question).

Consider a subsample of 10 people. In order to keep things simple, suppose each of them states the same WTP value to avoid 12 months of H – let us say £250 (which works out at about £10 for each fortnightly episode avoided). Suppose also that in response to Question E, 8 of these 10 would take a 1% risk of death in order to get a 99% chance of avoiding H for the rest of their lives, while one person would take a 2% risk and the remaining member of the subsample would only take a 0.1% chance. Averaging those responses gives an average QALY loss of 0.0101 per 12-month illness which, combined with the average WTP of £250 to avoid it, generates a value per QALY of just over £24,750.

Contrast this with the figure given by first combining each individual's WTP with their Question E response and *then* averaging. For eight respondents, the implied value of a QALY is $£250 \div 0.01$, which gives £25,000. For the respondent prepared to take a 2% risk, the figure is $£250 \div 0.02$, which gives £12,500. And for the respondent who will only accept a risk of 1 in 1000, the calculation is $£250 \div 0.001$, which produces £250,000. On this basis, the subsample mean is £46,250, i.e. almost double the figure arrived at by the other method.

The reason for the difference is this. Under the first method of computation, the two responses which differed from 0.01 were *added* to each other and to the eight 0.01s; and as they diverged from 0.01 in different directions but by almost the same amount, (+0.01 in one case, -0.009 in the other), this adding and averaging more or less cancelled them out, so that the subsample average diverged from 0.01 by just 0.0001. The net effect of this was to nudge the computed value of a QALY down to a little below £25,000.

However, under the second computation method, the (reciprocals of the) probabilities operate *multiplicatively* on the WTP responses before any averaging occurs. Thus the difference between 0.01 and 0.02 acts to halve the implied value of a QALY for that individual from £25,000 to £12,500 – a money difference of -£12,500. At the same time, the slightly smaller difference in the opposite direction between 0.01 and 0.001 acts to multiply by 10 the implied value – producing £250,000 as compared with £25,000 – a money difference

of +£225,000. When added and averaged at *this* stage, the two ‘outliers’ nowhere near cancel each other out: on the contrary, one individual’s implied value comes to more than all of the other nine put together, and this has the effect of almost doubling the mean.

So that is why the two methods of computation produce very different results; and the divergences in the example, striking though they may be, are dwarfed by the divergences in the actual sample, where still smaller probabilities had the effect of multiplying WTP responses by tens and hundreds of thousands, and even by millions.

So which is the *correct* method to use? In theory, if our analysis is based on the premise of each member of the population valuing the same QALY gain, it is the second method, i.e. the method used initially in the analysis of the survey data which generates stratospheric means. But, as with any theory, the validity and usefulness of the results depend crucially on the extent to which the underlying assumptions are valid. If we subscribe to the conventional precepts of welfare economics and if we could be confident that individuals have values and preferences which conform with standard assumptions and that their responses to our questions reveal those values with total accuracy and precision, the second method would be the appropriate one to use. In the light of the example which shows how one person’s value can outweigh the values of nine others, this may seem a surprising conclusion; but if all 10 really were reporting their true values accurately, and if the guiding principles entail giving each person and their values equal weight, then taking the mean of that distribution, however skewed, is the appropriate thing to do.

But what if certain assumptions do not hold perfectly? In particular, what if people are not always able to report their values with total accuracy, but are liable to give responses which contain elements of ‘noise’, error and/or bias? We are still some way from having very good models of the noise, error, bias and imprecision in human judgement; but we know enough to appreciate (a) that such things exist and (b) that they may not always be neutral or ‘white’ in their effects. So although we cannot say categorically how to model these factors, let us consider how they *might* impact upon the results of the present feasibility study.

Consider again the example set out a few paragraphs earlier. Suppose that all 10 respondents

not only have the same ‘true’ WTP to avoid the illness, but also would have their QoL diminished by the illness to exactly the same degree – that is, by 1%. If they each reported their values and judgements with complete accuracy, we should infer a mean value of a QALY of £25,000.

However, suppose that they do not all process probability judgements with unerring accuracy. To keep things simple, suppose that just two of them give erroneous responses – one reporting a willingness to accept a 2% risk, the other setting the ‘break-even’ risk of failure at 0.1%. Of course, this is only an illustration: we do not know enough about judgemental error to say exactly which and how many deviations from the underlying ‘true’ values are likely to occur. The point is, however, that if those two responses involve (seemingly small) errors, they can throw the estimates off in different ways and to very different degrees, depending on how they are processed. In the case of the first (and theoretically vulnerable) method of taking the arithmetic mean WTP separately from the arithmetic mean of the failure risks and then deriving a value of a QALY on the basis of the combination of these two means, the errors act to produce an estimate which is roughly £250 (and thus about 1%) below the true value; whereas in the case of the second (and theoretically preferable) method, the errors act multiplicatively before any averaging process occurs, and thereby produce an estimate which is £21,250 (and thus 85%) above the true value.

No general claim is being made here about the relative performance of the two methods: this is just a simple and convenient, albeit stylised, example. But what we do know is that people generally find probabilities quite challenging to manipulate, and their responses are liable to deviate a good deal, and in ways we do not well understand, from what is assumed by standard decision theories. (Recall the evidence relating to the sums of responses to Questions D and F, as reported in *Tables 35* and *36*.) So a method of calculation which is especially liable to magnify ‘errors’ by including them multiplicatively – and which we know in this case is liable to produce phenomenally large and implausible values – seems hard to justify as a basis for public policy.

Unfortunately, there is no way of knowing with any great confidence whether any serious distortions occur as the result of using the other method outlined here. It may be, as in the example, that this method has a tendency to understate the

'true' value. But equally, it is possible to produce examples where that is not the case; and even if it were the case in this study, we know too little about the nature of the imprecision/error in people's responses to say by how much any estimates diverge from the 'truth'. Perhaps the most that can be said is that this method, flawed though it may be, seems broadly compatible with the ways in which other values used by government departments have been derived; and to the extent that there is a desire to have a monetary value for a QALY established on a similar footing to the values used in DfT, DEFRA, etc., there appears to be some argument for this method of estimation, while keeping all the caveats in mind.

Concluding remarks

Speaking of caveats, it is important to reiterate that even if the data were unproblematic and the calculation of the monetary value of a QALY were straightforward, no value generated by *this* feasibility study could be regarded as a sound basis for policy because the sample used here was *not* representative of the population. The question is, then, whether the present study gives grounds for believing that (at least some of) the techniques explored here could be used in conjunction with a large and genuinely representative sample to produce a value robust enough to be used to guide policy, or whether the evidence suggests that no sufficiently robust value is likely (ever) to emerge.

There were a number of encouraging features. Interviewers generally found that respondents were engaged and interested, and very few interviews were aborted or resulted in large numbers of missing values. And even though many people in the UK are still somewhat resistant to the idea of paying for health care over and above tax contributions, the very small number of zero responses to B4 and especially B5 suggests that the 'protest' element was small and that people were willing to 'play the game'. Moreover, although the responses to B4 and B5 did not display the ideal 4:1 ratio that full sensitivity to duration might have elicited, there was a degree of sensitivity that *could* arguably have been compatible with the effect of budget constraints.

However, there are also a number of discouraging features. Although by no means new or unexpected, there was serious between-sample insensitivity to the size of the risk being eliminated in questions B6–B8. And although responses to

these questions should have been less affected by budget constraints and should therefore have shown greater sensitivity to duration, they did, if anything, exhibit rather less sensitivity to the 3-month:12-month difference; and also seemed to greatly underweight the 'remaining lifetime' duration relative to the 12-month scenario, with mean responses to B8 never more than 5 times higher than the corresponding means for B7, even though average remaining life expectancy was in the region of 30 years. It appeared that the use of probabilistic questions added complexity and dulled sensitivity.

Thus it may come as no surprise that some of the Part C questions, which used probabilities as their main 'currency', were also problematic. Although responses to Question D showed greater sensitivity to duration than was exhibited in the WTP questions, the mean response still appeared to greatly underweight remaining life expectancy relative to the next year. The summation of responses to Questions D and F showed a very substantial and seemingly systematic departure from the behaviour that would be consistent with standard assumptions.

It is hard to say whether the responses to Question E are 'reasonable' or not. Certainly, the distributions are heavily skewed: for both S and H the measure of skewness is greater than 2. Between a quarter and a third of respondents gave a 'failure risk' greater than the mean, while more than a third were unwilling to accept a risk of failure of 1 in 1000. Focusing on this latter statistic, does it really seem plausible that a condition involving significant disruption of at least some activities for between 8 hours and 3 days every couple of weeks entails a loss of welfare of less than 0.1%? Or is it that eliciting a response by means of a question involving the risk of immediate death induces ultraconservative responses from a significant minority of respondents? This is an open question. But if such responses are combined at the individual level with average WTP responses, they imply huge values of a QALY of the kind that generate the implausibly high means reported in this study.

On the other hand, if we compute means for B5 and for Question E separately and then use the ratio of these means to estimate the value of a QALY, the two figures obtained via different illness state descriptions are reasonably similar and not obviously outlandish.

Overall, then, this feasibility study sounds many notes of caution and points to a number of issues – particularly concerning the way respondents process probabilities and the extent to which their answers are sensitive to key dimensions – that would require further investigation before investing in a large representative national survey. Such

questions are potentially amenable to investigation using qualitative methods and entailing considerable cognitive testing during piloting, and it would be unwise to embark on a large-scale study to generate policy values without first undertaking extensive (and probably expensive) preparatory research of this kind.

Chapter 7

Conclusions

The main contributions of the research described in this report have been the development and initial application of two novel approaches to eliciting weights for QALYs along with one for estimating a monetary value of a QALY.

Detailed caveats have been listed within each of the preceding three chapters and, so, only broad research and policy conclusions will be offered here.

Weighting QALYs

The main results of the discrete choice and matching approaches show very different sets of weights for age and severity. In broad terms, the discrete choice results might be taken as suggesting that it is not worth weighting QALYs at all. By contrast, significant weights can be inferred from matching data, with gains for some groups being weighted up to 2.75 times higher than for others when using the more conservative aggregation method, the ratios going up to 4:1 using the less conservative method. Despite such differences in magnitude, discrete choice and matching weights did show similar patterns across age (but not severity) ranges.

There are two perspectives to take on the differences in results. First, it could be said that it would be premature to propose any particular set of QALY weights at this point in time: before that point is reached, there is scope for both further reconciliation and replication. Second, it might be argued that there is no scope for reconciliation and we need to choose between the results in light of the caveats of each.

Reconciliation and replication

With respect to the reconciliation of our findings it could, of course, be argued that it is not surprising that the discrete choice and matching methods led to different results. One factor has already been pointed out. The matching study involved holding health gains constant between the two options in any pairwise choice, and varying the

age and severity attributes, while the discrete choice questions presented a pair of scenarios with different health gains as well as levels of age and severity. It might be, therefore, that the matching procedure highlights age and severity while the discrete choice method dampens down their impact when the size of health gain is also varied. Indeed, if respondents have lexicographic preferences, whereby health gain matters above all else, it could be argued that the results of the discrete choice and matching approaches are entirely consistent. If this were the case, then the policy implications would require very careful thinking through and would pose major challenges to matching-based approaches in which choices presented hold QALY gains constant. Nevertheless, continuing to focus on the differences, there are additional limitations in the design of the discrete choice study which indicate caution around interpretation of results.

Other potentially important differences between the methods are listed in more detail in Appendix 9. Despite these, given that each respondent was asked a set of discrete choice questions and a set of matching questions, further detailed analysis may shed light on reasons for the differences.

Decision heuristics provides another possible avenue to reconciliation. First we have mentioned, in Chapter 5, that the lowest weight attached to the 60- to 80-year-olds in full health may have resulted from respondents thinking that 80% health is good enough for someone in such an age group. The lack of graphical representation of numbers of people in the matching study may have detracted respondents from the 'brutal' nature of trade-offs between persons, which, if dealt with otherwise, may have led to smaller trade-offs. With respect to the discrete choice approach, the simplest decision heuristic would be to compare the size of the areas representing the health gain, i.e. to maximize health and ignore age and severity. If each of these arguments had some validity, the 'true' result would be somewhere between the two.

It may be significant that the discrete choice and matching studies use different functional forms for age and severity weights.

In the discrete choice study, utility is modelled as a function of ‘age at onset’ (AO), ‘age at death if untreated’ (AD), ‘quality of life lost without treatment’ (QL), and ‘QALYs gained from treatment’ ($QALY$). Thus, the discrete choice functional form measures the health gain from an intervention by the size of the dark shaded area in *Figures 2 and 3*, i.e. it measures the health gain as the total number of QALYs. The functional form then treats this measure of ‘total QALY gain’ as one variable. The other three variables describe the lifetime health profile without treatment (the light shaded area). So, the weights produced by the discrete choice study are based on the properties of the light shaded area and the total area of the dark shaded area.

In contrast, the matching study focuses on the properties of the ‘dark shaded area’ of health gains. Recall that this study uses a ‘QALY grid’ of 20 cells, defined and numbered as in *Figure 16*. Each cell is assumed to have a subjective weight, which is applied to any health gains that occur in that cell. The subjective value of an intervention is modelled as the sum of the weighted health gains that are generated.

For example, consider the following case. In the absence of treatment, the lifetime health profile is given by $AO = 10$, $AD = 20$ and $QL = 0.7$ (i.e.

the patient is in full health to age 10, drops to 30% health from ages 10 to 20, and then dies). With treatment, the patient is maintained at 50% health from age 10 to age 30, and then dies. The health profile without treatment is represented by the light shaded area in *Figure 16*; the health gain from treatment is represented by the dark shaded area. The discrete choice study treats the effect of treatment as an undifferentiated gain of seven QALYs (i.e. an increase of 20 percentage points for 10 years, plus an increase of 50 percentage points for 10 years). The weight given to these QALYs is determined by the values of AO , AD and QL . In contrast, the matching study treats the effect of treatment as the creation of two QALYs in cell 2, one QALY in cell 5, two QALYs in cell 6 and one QALY in each of cells 9 and 10; each of these QALYs is given the weight of the cell in which it is located. Denoting the weight of each cell i by w_i , and the QALY gain in each cell i by q_i , a simple additive functional form can be specified as follows:

$$U = w_1q_1 + \dots + w_{20}q_{20}. \tag{6}$$

It is possible that the difference between the functional forms used in the two studies is responsible for the differences in their results. We tried to test this explanation by estimating equation (6) using the discrete choice data. One might expect that if the differences between the

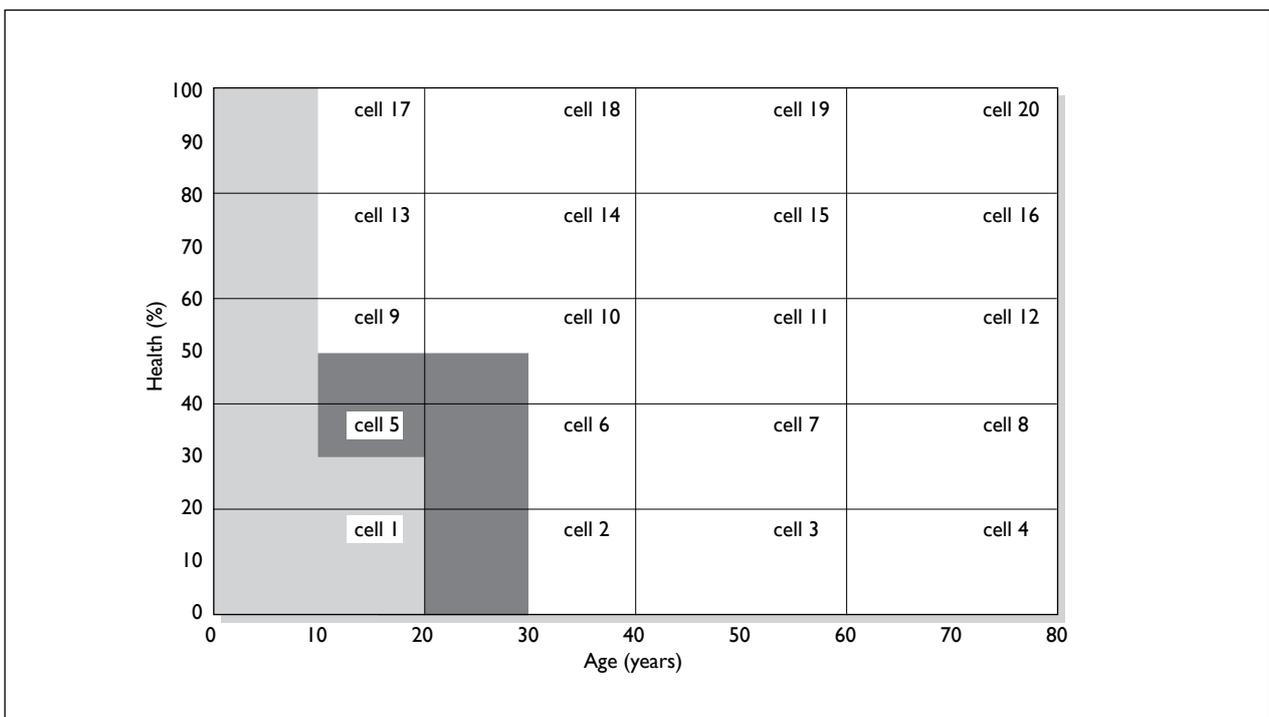


FIGURE 16 A health gain in the QALY grid.

results of the two studies were primarily due to the different functional forms used, this estimation exercise would generate weights similar to those found in the matching study. Alternatively, if it generated weights with some other systematic pattern, that might point towards some other explanation of the differences. (For example, if the weights were approximately equal in all cells, that would give further support to the conclusion that, when answering to discrete choice questions, respondents do not weight QALYs according to age or severity.) In fact, the estimated weights were highly variable, with no apparent pattern. The Bayesian information criterion clearly favoured the original model (described in Chapter 4, Predicted probability of choice approach) rather than equation (6). The lack of structure in these estimates is probably the result of econometric problems caused by multicollinearity in the data. The basic problem is that, in a typical discrete choice question, the health gains created by each intervention occur in cells that are adjacent to one another, creating strong positive correlation between the q_i values for adjacent cells. This reflects the fact that the questions used in the discrete choice study had been designed to estimate a very different model.

Procedurally, too, the exercises are very different (the results of the matching questions are derived from numbers being treated, which is always constant in the discrete choice) and discrete choice is perhaps a more complicated task (there are more things varying at any one moment, such as the health gain). Other such differences are outlined below. As pointed out in Chapter 1, these further differences represent concerns that each group within the Team, largely Newcastle-led and UEA-led, have with the piece of relativities work on which they did *not* lead. The Newcastle-led group think that both exercises have their merits, whereas the UEA-led group stands by the methods and results from the matching study.

Replication is also important. Aspects of each of these approaches are new and their applications to deriving relative weights for QALYs have generated significant challenges. Further survey work is required to inform the debate on weighting of QALYs with more confidence. With respect to discrete choice methods, as we have said, we would recommend pursuit of alternative experimental design strategies in order to address the issue of implausible scenarios while still maintaining desirable design properties, so ensuring we can

estimate the effects of interest with improved efficiency.

Fundamental differences between the approaches

It is possible, of course, to choose between the approaches in light of the caveats of each. Here, we outline those caveats without making such a choice. The discrete choice approach is new with respect to its application to deriving QALY relativities. This novelty could be viewed as advantageous. However, the approach raises some serious questions that would threaten the validity of the results. The main ones are:

- The issues raised by a compromised design by standards usually applied to discrete choice experimentation and whether the more pragmatic econometric approach overcomes these to any degree.
- The theoretical meaning that can be attached to the multiplicative functional form that underlies the analysis of the discrete choice data. More specifically, the meaning of multiplying QALYs by two age variables as well as by severity, and the challenges of multicollinearity such a model engenders. In defence of this model, however, it is not clear what form such a function should take and the use of squared and cubic terms might reasonably have been expected to pick up any non-linearities implied by the functional form arising from the matching data. Also, as pointed out in Chapter 4, although, it may appear that a multiplicative model of the form $QALYs \times AGE \times SEVERITY$, with just one age-related variable, would make more theoretical sense, we took the more pragmatic view that this would leave too much riding on the 'age at onset' variable in terms of explaining what respondents might be thinking about in relation to age, and so we included age at death as well. In addition, this was the best performing model empirically, which, it could be argued, is important for estimating weights.
- Whether the transformations of original variables to a QALY composite and QoL lost represent too much of a distortion from the variables as presented to respondents in the survey.
- The challenges of coping with severity as described in Chapter 4, Further investigation of severity.

The matching approach builds on a method used previously in the literature. The results are not out of line with such earlier studies, the pattern of weights being consistent across aggregation methods. Nevertheless, in addition to QALY gains being held constant in the matching questions, there are some issues with the matching that remain to be resolved. These are that:

- There is an underlying assumption that numbers of beneficiaries presented in matching scenarios can be multiplied by the number of QALYs (implicitly) presented, and it is not clear if respondents were thinking in such a manner.
- While the ‘whole column’ and ‘whole row’ tests looked at the issue of larger QALY gains, it is unclear whether we can generalise from weights obtained by offering four QALY gains to gains of fewer than four QALYs and it is acknowledged that many interventions yield a much smaller benefit than four QALYs.
- Although a general pattern has been detected, the appropriate magnitude of the weights remains to be established. This is not a problem of the matching methodology per se, but rather an inevitable part of any elicitation exercise which is used to guide policy.

In summary, the extent to which either approach yields results that are entirely consistent with social preferences is uncertain. This work has indicated some important patterns in the data. It could be argued that the old and the healthy should receive lower weights. The challenges are with the actual numbers to adopt and so further validation and testing would be required before implementation in policy.

The feasibility of estimating a monetary value of a QALY

It was never the intention in the valuation study to produce a monetary value of a QALY for use in policy deliberations. The purpose was simply to assess the feasibility of estimating such a value.

There is scope to use the current data set to further explore different ways of combining WTP and SG data to arrive at a value of a QALY, and to examine the sensitivity of results to different assumptions about noise in the data – and especially in the probability responses generated by the SG questions. Further consideration needs to be given

to aggregation issues: for example, the weight to attach to means versus medians, and whether to use means of individual WTP/SG combinations or combinations of mean WTP and mean SG values.

The feasibility study has demonstrated that there are considerable challenges involved in trying to elicit a robust monetary value of a QALY. However, the low number of protest responses demonstrates that respondents are comfortable with these types of question. Any future national sample survey should be preceded by further extensive qualitative research and cognitive testing to resolve the main questions identified in the present study.

Implications for research and practice

Implications for practice

Given the methodological nature of the research reported, implications for practice are limited, but twofold:

- On relativities, it could be said that it would be premature to propose any particular set of QALY weights at this point in time: before that point is reached, there is scope for both further reconciliation and replication. However, it might equally be argued that there is no scope for reconciliation and we need to choose between the results in light of the caveats of each.
- On valuation, it was never the intention to conduct a representative survey using a definitive method. The main recommendation, therefore, is that any future national sample survey should be preceded by further extensive qualitative research and cognitive testing to resolve the main questions identified in the present study.

Research recommendations

The research recommendations arising from the study are that:

- The findings from the relativities study indicate that more work is required in the short term to reconcile the results obtained, although fundamental differences between the methods and results reported may challenge such reconciliation.
- Longer term, and still with respect to relativities, further methodological research should attempt to account for some of the

deficiencies of the methods (especially the particular discrete choice approach used in SVQ).

- Building on the results of the innovative methods that have been devised in this study to derive relative weights, further replication of these results is required to address this important policy issue.
- With respect to valuation, shorter-term work is required around the issues of aggregation, combining WTP and SG values and the appropriateness of different measures of central tendency.
- Longer term, more qualitative and cognitive research is required around two valuation issues in particular: first, the problem of identifying health states to present to respondents which are 'minor enough' for people to be able to express their WTP but not so minor that respondents will accept only minuscule risks of death when responding to SG type questions; and, related to the first, the extent to which 'noise' and 'error' in people's responses might generate extreme and unreliable figures.



Acknowledgements

The authors would like to acknowledge the support of the (then) National Coordinating Centre for Research Methodology of the Department of Health, as well as the National Centre for Social Research, who conducted the two surveys reported on in this monograph.

Contribution of authors

Rachel Baker contributed to the design of the study, analysis and interpretation of data, and drafting and editing of the report. She led the Q methodology and qualitative analysis and drafted Chapter 3. Ian Bateman contributed to the conception and design of the study, analysis and interpretation of data, revision of the report and approval of the final version. Cam Donaldson (lead author and principal investigator) contributed to the conception and design of the study, analysis and interpretation of data, and drafting and editing of the report. Michael Jones-Lee, Jose

Luis Pinto Prades, Mandy Ryan, Phil Shackley, Richard Smith and Robert Sugden contributed to the conception and design of the study, analysis and interpretation of data, and revision and editing of the report. Emily Lancsar and John Wildman contributed to the design of the study, analysis and interpretation of discrete choice data, and drafting and revision of the report. Graham Loomes contributed to the conception and design of the study, analysis and interpretation of data, and drafting and editing of the report. He led the valuation study and drafted Chapter 6. Helen Mason contributed to the design of the study, analysis and interpretation of data, and revision and editing of the report. Maria Odejar contributed to the analysis and interpretation of discrete choice data. Angela Robinson contributed to the conception and design of the study, analysis and interpretation of data, and drafting and editing of the report. She led the analysis of person trade-off relativities study data and drafted Chapter 5.



References

1. Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ* 1986;**5**:1–30.
2. Williams A. Economics of coronary artery bypass grafting. *Br Med J* 1985;**291**:326–9.
3. National Institute for Health and Clinical Excellence. *Guide to the methods of technology appraisal*. London: NICE; 2007.
4. Donaldson C, Atkinson A, Bond J, Wright I. Should QALYs be programme-specific? *J Health Econ* 1988;**7**:239–57.
5. Weinstein M. A QALY is a QALY is a QALY – or is it? *J Health Econ* 1988;**7**:489–91.
6. Williams A. QALYs and ethics: a health economist's perspective. *Soc Sci Med* 1996;**43**:1795–804.
7. Rawlins M, Culyer A. National Institute for Clinical Excellence and its value judgements. *BMJ* 2004;**329**:224–7.
8. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence their decisions? A binary choice analysis. *Health Econ* 2004;**13**:437–52.
9. Loomes G. Valuing life years and QALYs: transferability and convertibility of values across the UK public sector. In: Towse A, Pritchard C, Devlin N, editors. *Cost effectiveness thresholds: economic and ethical issues*. London: King's Fund and Office of Health Economics; 2002.
10. House of Commons Health Select Committee. *National Institute for Health and Clinical Excellence, first report session 2007–08, volume 1*. London: The Stationery Office; 2007.
11. Nord E, Street A, Richardson J, Kuhse H, Singer P. The significance of age and duration of effect in social evaluation of health care. *Health Care Anal* 1996;**4**:103–11.
12. Nord E. Towards cost–value analysis in health care? *Health Care Anal* 1999;**7**:167–75.
13. Dolan P, Olsen JA. *Distributing health care: economics and ethical issues*. Oxford: Oxford University Press; 2002.
14. Cropper ML, Aydede SK, Portney PR. Preferences for life saving programs: how the public discounts time and age. *J Risk Uncert* 1994;**8**:243–65.
15. Johannesson M, Johannsson P-O. The economics of ageing: on the attitude of Swedish people to the distribution of health care resources between the young and the old. *Health Policy* 1996;**37**:153–61.
16. Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. *Health Econ* 2005;**14**:197–208.
17. Schwappach DL. Resource allocation, social values and the QALY: a review of the debate and empirical evidence. *Health Expect* 2002;**5**:210–22.
18. Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. *Health Econ* 2000;**9**:137–48.
19. Schwappach DL. Does it matter who you are or what you gain? An experimental study of preferences for resource allocation. *Health Econ* 2003;**12**:255–67.
20. Johri M, Dmaschroder IJ, Zikmund-Fisher BJ, Ubel PA. The importance of age in allocating health care resources: does intervention-type matter? *Health Econ* 2005;**10**:461–74.
21. Mason H. *Monetary valuation of health outcomes for use in national policy formulation*. Newcastle upon Tyne: PhD; 2007.
22. Dupuit J. On the measurement of utility of public works. [orig. 1844]. *Int Econ Papers* 1952;**2**:83–110.
23. Davis R. Recreation planning as an economic problem. *Nat Resource J* 1963;**3**:239–49.
24. Jones-Lee M. *The economics of safety and physical risk*. Oxford: Blackwell; 1989.
25. Gafni A. Willingness to pay as a measure of benefits: relevant questions in the context of public decision making about health care programmes. *Med Care* 1991;**29**:1246–52.
26. O'Brien B, Gafni A. When do the 'dollars' make sense? Toward a conceptual framework for contingent valuation studies in health care. *Med Decis Making* 1996;**16**:288–99.

27. Acton JP. *Evaluating public programmes to save lives: the case of heart attacks*. Report No.: R950RC. Santa Monica: RAND Corporation; 1976.
28. Donaldson C. Valuing the benefits of publicly-provided health care: does 'ability to pay' preclude the use of 'willingness to pay'? *Soc Sci Med* 1999;**49**:551–63.
29. Donaldson C, Birch S, Gafni A. The pervasiveness of the 'distribution problem' in economic evaluation in health care. *Health Econ* 2002;**11**:55–70.
30. Johannesson M, Jonsson B, Borgquist L. Willingness to pay for anti-hypertensive therapy – results of a Swedish pilot study. *J Health Econ* 1991;**10**:461–74.
31. Olsen JA, Smith R. Theory versus practice: a review of 'willingness to pay' in health and health care. *Health Econ* 2001;**10**:39–52.
32. Smith RD. Construction of the contingent valuation market in health care: a critical assessment. *Health Econ* 2003;**12**:609–28.
33. Donaldson C. Eliciting patients' values by use of 'willingness to pay': letting the theory drive the method. *Health Expect* 2001;**4**:180–8.
34. Olsen JA, Donaldson C, Pereira J. The insensitivity of 'willingness to pay' to the size of the good: new evidence for health care. *J Econ Psychol* 2004;**25**:445–60.
35. Olsen JA, Donaldson C, Shackley P, Group E. Implicit versus explicit ranking: on inferring ordinal preferences for health care programmes based on differences in willingness-to-pay. *J Health Econ* 2005;**24**:990–6.
36. Olsen JA, Kidholm K, Donaldson C, Shackley P. Willingness to pay for public health care: a comparison of two approaches. *Health Policy* 2004;**70**:217–28.
37. Yeung RYT, Smith RD, McGhee SM. Willingness to pay and size of health benefit: an integrated model to test for 'sensitivity to scale'. *Health Econ* 2003;**12**:791–6.
38. Smith RD. Sensitivity to scale in contingent valuation: the importance of the budget constraint. *J Health Econ* 2005;**24**:515–29.
39. Protière C, Donaldson C, Luchini S, Moatti JP, Shackley P. The impact of information on non-health attributes on willingness to pay for multiple health care programmes. *Soc Sci Med* 2004;**58**:1257–69.
40. van Exel NJA, Brouwer WBF, van den Berg B, Koopmanschap MA. With a little help from an anchor: evidence of starting point bias in contingent valuation of informal caregiver time inputs. *J Socio-Econ* 2006;**35**:836–53.
41. Carthy T, Chilton S, Covey J, Hopkins L, Jones-Lee M, Loomes G, *et al*. On the contingent valuation of safety and the safety of contingent valuation: Part 2 – the CV/SG 'chained' approach. *J Risk Uncert* 1999;**17**:187–213.
42. Johannesson M. The relationship between cost-effectiveness analysis and cost-benefit analysis. *Soc Sci Med* 1995;**41**:483–9.
43. Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. *J Health Econ* 1997;**16**:1–31.
44. Culyer A, McCabe C, Briggs A, Claxton K, Buxton M, Akehurst R, *et al*. Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. *J Health Serv Res Policy* 2007;**12**:56–8.
45. Bate A, Murtagh M, Donaldson C. Managing to manage scarce resources in the English NHS. What can economics teach? *Health Policy* 2007;**84**:249–61.
46. Hirth RA, Chernew ME, Miller E, Fendrick M, Weissert WG. Willingness to pay for a quality-adjusted life year: in search of a standard. *Med Decis Making* 2000;**20**:332–42.
47. Gyrd-Hansen D. Willingness to pay for a QALY. *Health Econ* 2003;**12**:1049–60.
48. Mason H, Marshall A, Donaldson C, Jones-Lee M. *Estimating a willingness to pay based value of a QALY from existing UK values of prevented fatalities and serious injuries*. Birmingham: National Coordinating Centre for Research Methodology; 2005.
49. Byrne MM, O'Malley K, Suarez-Almazor ME. Willingness to pay per quality adjusted life year in a study of knee osteoarthritis. *Med Decis Making* 2005;**25**:655–66.
50. King JT Jr, Tsevat J, Lave JR, Roberts M. Willingness to pay for a quality adjusted life year: implications for societal health care resource allocation. *Med Decis Making* 2005;**25**:667–77.
51. Coast J, Horrocks S. Developing attributes and levels for discrete choice experiments using qualitative methods. *J Health Serv Res Policy* 2007;**12**:25–30.

52. Stephenson W. *The study of behavior: Q-technique and its methodology*. Chicago, IL: University of Chicago Press; 1953.
53. Brown SR. *Political subjectivity: applications of Q methodology in political science*. London: Yale University Press; 1980.
54. Baker R, Thompson C, Mannion R. Q methodology in health economics. *J Health Serv Res Policy* 2006;**11**:38–45.
55. Brown SR. A primer on Q methodology. *Operant Subj* 1993;**16**:91–138.
56. Greene WH. *Econometric analysis*. 6th edition. New Jersey: Prentice-Hall; 2007.
57. Small K, Rosen H. Applied welfare economics with discrete choice models. *Econometrica* 1981;**49**: 105–30.
58. Lancsar E, Savage E. Deriving welfare measures from discrete choice experiments: inconsistency between current methods and random utility and welfare theory. *Health Econ* 2004;**13**:901–7.
59. Dolan P, Green C. Using the person trade-off approach to examine differences between individual and social values. *Health Econ* 1998;**7**:307–12.
60. Chilton S, Covey J, Hopkins L, Jones-Lee M, Loomes G, Pidgeon N, *et al*. Public perceptions of risk and preference-based values of safety. *J Risk Uncert* 2002;**25**:211–32.
61. Johannesson M, Johannsson P-O. Is the valuation of a QALY gained independent of age? Some empirical evidence. *J Health Econ* 1997;**16**:589–99.
62. Busschbach JJV, Hensing DJ, de Charro FT. The utility of health at different stages in life: a quantitative approach. *Soc Sci Med* 1993;**37**:153–8.
63. Ubel PA, Spranca MD, DeKay ML, Hershey JC, Asch DA. Public preferences for prevention versus cure: what if an ounce of prevention is worth only an ounce of cure? *Med Decis Making* 1998;**18**:141–8.
64. Dolan P, Tsuchiya A. Health priorities and public preferences: the relative importance of past health experience and future health prospects. *J Health Econ* 2005;**24**:703–14.

Appendix I

Focus group introduction

Introductions

[Introduce ourselves]

This study is sponsored by the Department of Health (DoH) whose role is to ensure that the money available for health care is spent in the best possible way. Not all needs can be met from available resources. There are many different ways in which money can be spent in improving our health, but the DoH sometimes has to prioritise between these in deciding which are the most important.

The DoH wants some guidance from members of the public like yourselves about whether more resources should be put towards some sorts of health-care activities than others, which may benefit different kinds of people. The DoH is more interested in the general principles the public wish to be taken into consideration when prioritising health care, rather than their views about particular treatments or drugs. For example, some health interventions save lives whilst others improve the quality of people's lives. Some interventions cure people of a disease or illness whilst others prevent disease or illness from occurring in the first place. And, of course, interventions may benefit different types of people.

This survey has been designed to find out what sorts of things might or might not matter to people. There are no right or wrong answers to this: we just want to know what you think.

[Tell them we are taping and start tape]

[Ask them each to say their first name and how they travelled to the group to get their voices on tape]

Appendix 2

Factor loadings

ID	Sex	Age	Factor loadings		
			F1	F2	F3
1	m	44	0.23	0.44x	0.37
2	f	52	0.07	0.56x	0.43
3	m	84	0.51	0.12	0.52
4	m	59	0.36	0.40x	0.14
5	m	58	0.43	0.42	0.20
6	f	52	-0.29	-0.57x	0.21
7	m	63	0.28	0.35	0.65x
8	f	56	0.60x	0.30	0.43
9	f	42	0.24	0.28	0.27
10	m	48	0.75x	0.04	0.16
11	m	36	0.35	0.33	0.53x
12	m	39	0.60x	0.14	0.32
13	m	53	0.36	0.04	0.47x
14	f	40	0.73x	0.36	0.10
15	m	20	0.50x	0.20	0.42
16	m	58	0.10	0.84x	0.07
17	f	41	0.73x	0.30	0.26
18	f	50	0.55x	0.45	0.22
19	m	25	0.12	0.29x	0.26
20	f	22	0.65x	0.27	0.18
21	f	42	0.68x	-0.02	0.34
22	f	26	0.36x	-0.01	0.23
23	m	37	0.52	0.52	0.16
24	m	28	0.44	0.11	0.58x
25	f	45	0.24	0.27	0.57x
26	m	35	0.19	0.24	0.35x
27	m	27	0.21	0.00	0.66x
% explained variance			20	12	14

Significant loadings are shown in bold type. The significance level for factor loadings is taken as 2.5 (SE). SE represents standard error which is defined as $1/\sqrt{N}$ where N is the number of statements in the Q set. In this case then, $2.5 (SE) = 2.5 (1/\sqrt{46}) = 0.37$.
 Defining sorts are identified by x. The automatic flagging procedure in PQ METHOD software was used to identify defining sorts.

Appendix 3

Q sort statements and factor scores

	Statement	Factors		
		F1	F2	F3
1	Whether or not an illness could have been avoided needs to be taken into account. The cause of the illness matters; more priority should be given to things arising from environmental causes and allergies.	0	-3	-1
2	Whether or not people have caused the illness themselves should not be relevant. If someone has got an illness through smoking, they are just as worthy of treatment as someone else. It's wrong for the health service to make moral judgements about people's lifestyles.	1	3	-1*
3	People who live a healthy lifestyle should be prioritised because they would respond better to treatment.	-5	-1*	-4
4	You can't prioritise health care on the basis of people's lifestyle – like not giving equal priority to smokers – where would you draw the line? Obesity? People who do dangerous sports? You just have to treat everyone who is ill the same regardless of lifestyle.	3*	0	0
5	If some people have lung cancer through no fault of their own (e.g. because of their work, or environment) they should be given higher priority than people who get lung cancer because they smoke.	-3	-2	0
6	If someone is given treatment, like George Best, and then abuses their treatment, they should not be given repeated chances. If there are finite resources and a person has failed to take advantage of it, someone else should be given a chance.	4*	1	1
7	People don't always have control over their lifestyles; it can be rooted in their background which causes problems of addiction or other things. It's not as simple as people choosing to damage their health.	1	0	0
8	Illnesses which have a negative impact on other people (e.g. the patients' families) should be given priority.	-2	1	-2
9	People who smoke and drink pay enough in extra taxes to pay for their own health care.	-2	-4	-5*
10	Sometimes you have to make decisions and sometimes you have to give priority towards younger people because improving health for an 80-odd year-old might not make a difference to their lives.	-2	2*	0
11	Life is equally valuable whether you are young or old.	5*	1*	3*
12	If people have had some treatment in the past for the same problem which hasn't resolved the problem, then they should be given priority for treatments.	0	1	-2*
13	Age shouldn't come into it, unless you're talking about children. Children's health should be given priority over adults.	0	0	4*
14	The age of the patient is important; if you were treating children rather than older people then you would have a longer improved life.	-4*	2*	-1*
15	Everybody, no matter what you are, whether you are young or old, should get the same access to and choice of treatment.	5*	-1*	3*
16	You should prioritise the younger age group, because they are still able to have children.	-5*	-1	-3
17	For relatively minor conditions patients who are of working age should take priority over people who are retired.	-3	1*	-4

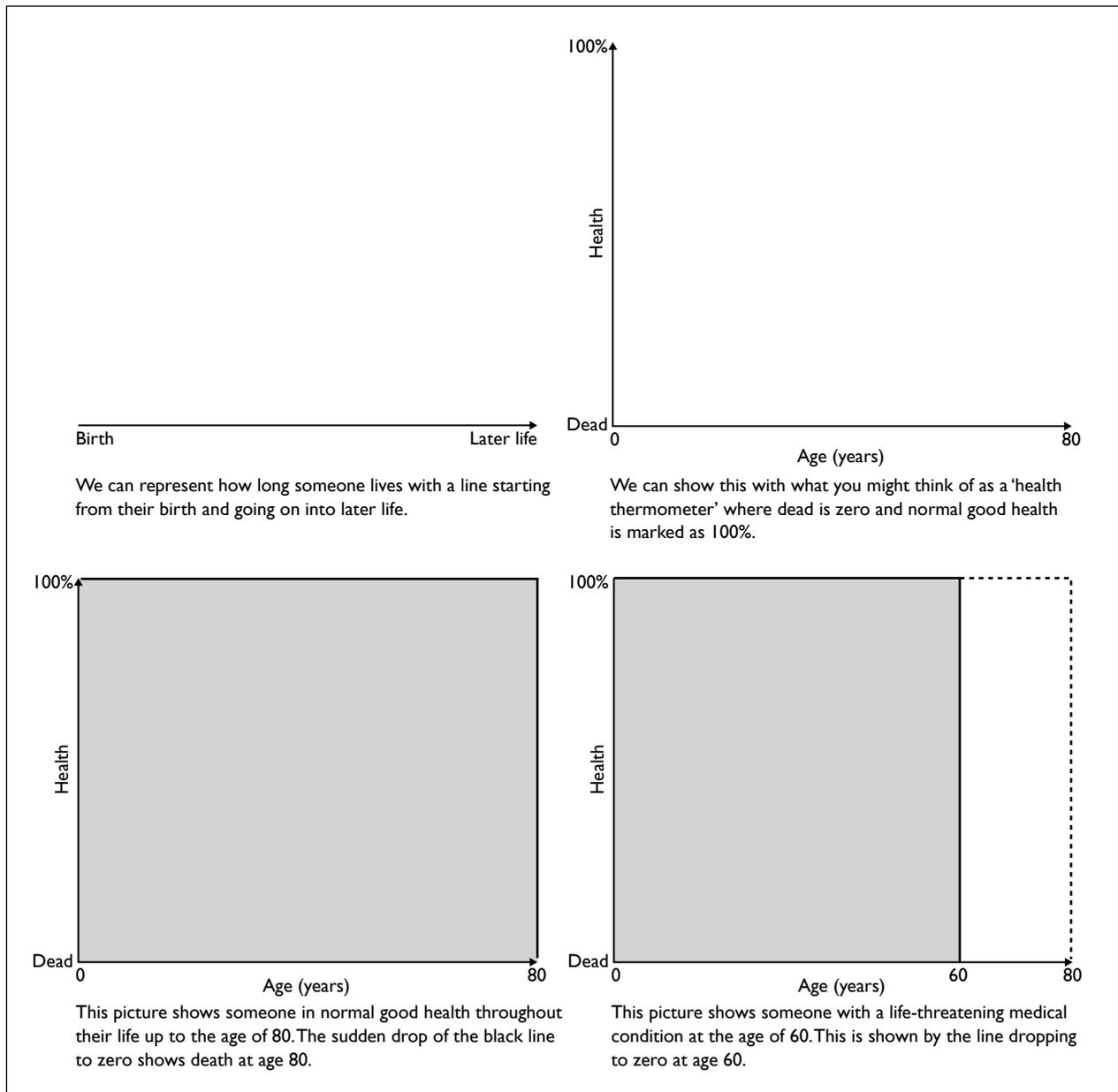
	Statement	Factors		
		F1	F2	F3
18	There should be 'positive discrimination' towards people who are disadvantaged and in ill health because they've got a lot to contend with already.	0	-4*	-1
19	The important thing is that people can fulfil their 'role' in society, whatever that is – work, caring for people or housework.	2	2	1
20	People with dependants should be prioritised over people without dependants because their treatments would benefit others as well as the patients themselves.	-4	2*	-4
21	Older people deserve to be given priority. They have been paying into the NHS all their lives; they deserve to be able to draw on those resources when they need it.	-1	-4*	-2
22	It's important to take into account how many people stand to benefit from a treatment. If the same health benefit could be given either to 10 people – each getting a large benefit – or 1000 people – each getting a small benefit – I would want to treat 1000.	-1	-1	2*
23	Priority should be given to preventive health care especially education in schools about diet and lifestyle choices.	3	4	5
24	Whether or not people are currently working should be taken into account when we prioritise health care.	-4	-4	-4
25	People with dependants should not be given priority over people without dependants. A human life is a human life; I think it should be irrelevant how many dependants they've got.	4*	-2*	2
26	Poorer people should be given priority because they don't have the same opportunities to take care of their own health.	-1	-5*	-3
27	People choose to have children – they shouldn't be prioritised over people who don't have children. Why should people be penalised for not having children?	2*	-2*	1*
28	Whether or not patients can contribute financially towards the cost of the treatment should be taken into account because it would allow you to treat more people who can't afford to 'go private'.	-3*	-5*	-1*
29	The quality of life of patients and their life expectancy are the most important things. The characteristics of patients like whether they are employed, or whether they have dependants, or what gender they are shouldn't matter.	3	5*	2
30	Social class should make no difference whatsoever for prioritising health care. If people need treatment, they need treatment. How well off they are shouldn't come into it.	4	4	4
31	The amount of health and quality of life improvement is the most important. It's about getting the greatest benefit for the most people.	2*	4	4
32	Priority should be given to people who can't take care of themselves and their own basic needs.	2	-3*	3
33	New treatments should be targeted towards diseases for which there is currently no other treatments available. Even if they are only of limited benefit, it's important that people get <i>something</i> .	-1*	-3*	2*
34	It is more important to save one life than it is to improve the lives of many by only a small amount.	-2	-1	1
35	The decisions about which services to fund, and how to spend NHS money should be made by a range of experts with a lot of information and experience of the issues, not the general public.	-3	0*	3*
36	The quality of life of patients before treatment is the most important thing. Priority should be given to people whose quality of life is the lowest, even if we can only improve it by a small amount.	0	-3*	1

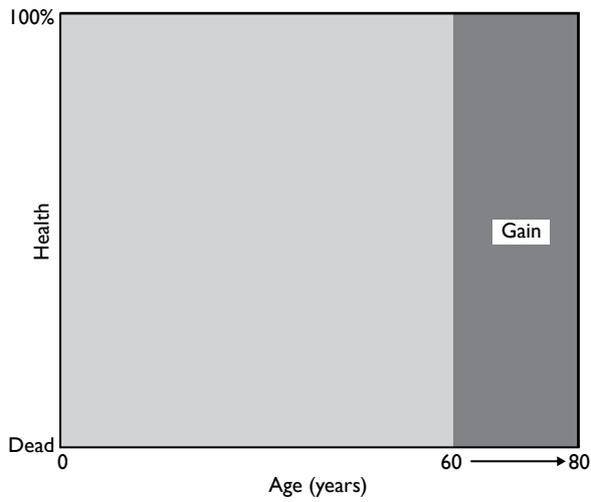
	Statement	Factors		
		F1	F2	F3
37	You can't just say age matters or doesn't matter. There are so many other things that you would want to know about the treatment. It's complex and you can't take one thing in isolation.	2	3	2
38	Priority should be given to treatments without side effects. It's important that health care does no harm to people.	1	2	0
39	Treatments which add years to life are more important than treatments which only have an impact on quality of life.	-1	-2	-2
40	Priority should be given to preventive health care rather than always focusing on cure once people are ill.	3*	5	5
41	Health care should be based on <i>need</i> , not on social circumstances, or addiction or weight or smoking or anything else.	4	3	4
42	If you're prioritising health services it should just be based on the principle that a person is in a queue and it's their turn, not anything about the people themselves. That's the fair way.	0	-2	-3
43	I would want to consider how stressful a treatment is and prioritise treatments which would 'disturb' the patients less, which were less invasive.	1	3	-2*
44	It's no good saving lives if the quality of those lives is really bad. Some treatments are keeping people alive for too long. You've got to have a decent quality of life otherwise what's the point of being alive.	1*	4*	-3*
45	It's important to consider whether or not someone has dependants when you're prioritising health care, because if something happens to them then something will have to be done for their dependants.	-2*	0	0
46	People who have already benefited from a lot of health care should take second place to people who have not used the health service as much.	-4	0*	-5

Note: * denotes those statements which distinguish one factor from the other two factors ($p < 0.01$). Shaded statements are consensus statements which do not distinguish between any two pairs of factors (non-significant at $p > 0.01$).

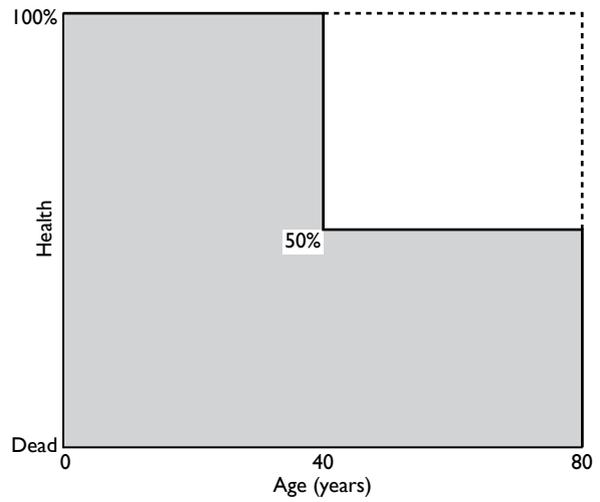
Appendix 4

Diagrammatic approach to explaining QALYs

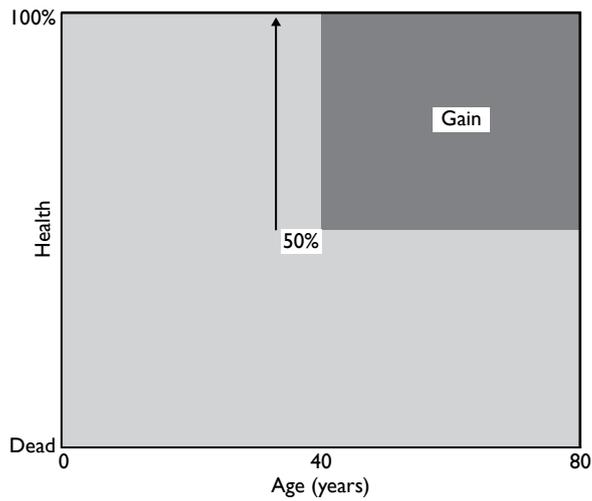




This picture shows a successful treatment for that condition. The treatment gives them an extra 20 years' life expectancy.



Some medical conditions do not affect life expectancy. This picture shows someone aged 40 who suffers a condition which reduces their health from 100% to 50% for the rest of their life.



With successful treatment their health can be restored to 100%.

Appendix 5

Constraints on DCE design

1. Age at death must be no less than age of patients.
2. The sum of expected age at death if untreated and gain in life expectancy must not exceed 80 years.
3. If age at death if untreated is equal to age of patients, and gain in quality of life with treatment is 0, then gain in life expectancy must be 0.
4. If age at death if untreated is equal to age of patients, then quality of life if untreated is 0. If age at death if untreated is greater than age of patients, then quality of life if untreated is not 0.
5. Since the gain in quality of life with treatment has a maximum of 100%, the sum of quality of life if untreated and gain in quality of life with treatment is less than or equal to 100%.
6. Gain in quality of life with treatment and gain in life expectancy cannot both be 0.
7. If age at death if untreated is equal to age of patients, and gain in life expectancy is 0, then gain in quality of life with treatment is 0.
8. If gain in life expectancy is not 0, then the sum of quality of life if untreated and gain in quality of life with treatment cannot be 0.

Appendix 6

Version 2 of the willingness-to-pay/ standard gamble questionnaire

P2487

CARD NO.	1009 – 1010
BATCH NO.	1011 - 1015
VERSION NO.	1016 - 1017

VALUING HEALTH:

INDIVIDUAL QUESTIONNAIRE

Version 2

RESPONDENTS SERIAL NUMBER: *1001 – 1006* *1007* *1008*

Checkletter Person no

RESPONDENTS FIRST NAME: _____

INTERVIEWER NUMBER:
1018 - 1021

START TIME OF INTERVIEW (USE 24 HR CLOCK):
Hours Mins
1022-1023 1024-1025

1026-1035 Spare

INTRODUCTION*For
office
use
only***INTERVIEWER READ OUT:**

Thank you for taking part in this research, which is intended to inform health policymakers about the values and preferences of members of the public.

In the course of the interview, we will ask questions about some specific states of health. However, these are just examples and we are not so much concerned with particular illnesses as with health in the more general sense. So when we ask you to consider particular examples, we'd like you to think about how such illnesses might affect your ability to undertake your usual activities, as well as how any pain and discomfort might affect your enjoyment of life.

We will be interviewing a cross-section of the public, so we are asking you to answer just on your own behalf. If you do that and everyone else does that, we will get a good overall picture. All the answers you give are confidential and will only be seen by members of the research team.

First, I'd like to start by asking you to complete some questions by yourself.

INTERVIEWER: HAND PART A TO RESPONDENT.

PART A: (Respondent to complete)

Let us start by considering your current state of health. For each section in turn, please circle the number of whichever statement best describes your own situation **today**.

QA1	Mobility: (How easily you get around by yourself)	I have no real problems in walking about	1	1036
		I have occasional minor problems in walking about	2	
		I have frequent but mostly minor problems in walking about	3	
		I have quite a lot of difficulties in walking about	4	
		I cannot get about much at all	5	
QA2	Self Care: (Washing, dressing, feeding, etc.)	I have no real problems with self-care	1	1037
		I have occasional minor problems with self-care	2	
		I have frequent but mostly minor problems with self-care	3	
		I have quite a lot of difficulties with self-care	4	
		I cannot manage much self-care at all	5	
QA3	Usual activities: (Work, study, housework, family/hobby/leisure activities)	I have no real problems with such activities	1	1038
		I have occasional minor problems with such activities	2	
		I have frequent but mostly minor problems with such of activities	3	
		I have quite a lot of difficulties with such activities	4	
		I cannot manage many of these sorts of activities at all	5	
QA4	Pain and discomfort:	I have no real problems with pain or discomfort	1	1039
		I have occasional minor pain and discomfort	2	
		I have frequent but mostly minor problems with pain and discomfort	3	
		I have quite a lot of pain and discomfort for most of the time	4	
		I am in a lot of pain and discomfort all or nearly all of the time	5	
QA5	Anxiety or Depression:	I have no real problems with anxiety or depression	1	1040
		I have occasional minor problems with anxiety or depression	2	
		I have frequent but mostly minor episodes of anxiety or depression	3	
		I have quite a lot of difficulties with anxiety or depression	4	
		I suffer from very serious anxiety or depression	5	

QA6

Now we would like you to tell us how good or bad you consider your health to be by putting a cross on the 'health thermometer' on the right.

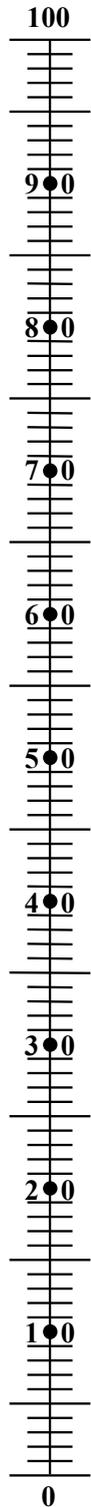
On this thermometer, 100 represents good health **FOR SOMEONE OF YOUR AGE**. Please think about good health for someone of your age, rather than the best health imaginable for any age.

At the bottom, zero represents a state of health you regard as being as bad as being dead.

What number do you think represents your overall current state of health as compared with good health for someone of your age?

Thank you. Please hand this booklet back to the interviewer

As good as it could be for someone of your age



As bad as being dead

PART B: (Interviewers to administer)

B1 RECORD THE NUMBER THE RESPONDENT INDICATED AT QA6:

IF NOT CLEAR WHICH NUMBER, ASK RESPONDENT TO CLARIFY

--	--	--

1041-
1043

B2i GIVE HEALTH CARDS A, E, M and V FACE-UP TO RESPONDENT

Here are some cards which describe various episodes of ill health. Please read through these four cards, and think about how bad **you personally** think you would find them, compared with one another. Then set them out in order from the one you personally consider **least bad** down to the one you consider **worst**.

B2ii AFTER RESPONDENT HAS SORTED THE FIRST FOUR HEALTH CARDS, GIVE CARDS S & T TO RESPONDENT

Here are two more descriptions. How do you think they compare with the ones you have just sorted?

Please add them to the other four, so that you end up with all six arranged in order so that the **least bad** one is at the top and then they get worse as you go down until the one at the bottom is the one you consider **worst** of all.

B2iii AFTER RESPONDENT HAS SORTED CARDS S AND T, GIVE HEALTH CARD D TO RESPONDENT

This card is sudden but painless death. Do you think any of the illnesses described on those cards is worse than being dead?

If so, put the 'sudden death' card **above** any of the illnesses that you consider worse than being dead. But if you think that even the worst of those illnesses is better than being dead, put the sudden death card right at the bottom to show that you think it is worst of all.

B2A AFTER RESPONDENT HAS SORTED ALL CARDS, TURN CARDS OVER AND RECORD THE ORDER:

	<i>Record letter</i>	
Least bad (top)		1044
2 nd		1045
3 rd		1046
4 th		1047
5 th		1048
6 th		1049
Worst (bottom)		1050

PART B: (Interviewers to administer)									
B3	<p>READ OUT: Now I'd like to know how valuable it would be to you to avoid some of the illnesses you have just looked at. There are various ways of asking these sorts of questions. One way is to ask people what money value they would place on different treatments. The money value you put on different things is a good way of comparing how important they are to you. The amount you would be willing to pay for some treatments if you had to gives an indication of how you value health gains from those treatments compared with other things you might want to spend your money on. Let's start with a 'practice' question.</p>								
B3 a	<p>READ OUT: Imagine you are on holiday abroad in quite a remote place and you get a bad sore throat – bad enough that it is painful every time you swallow food. If you don't take anything for it, suppose it will last for 3 days. However, someone can arrange to have medicine delivered to you that will cure your throat infection within 24 hours, thereby saving you 2 of the 3 days of painful sore throat. Think about what would be the MOST you would be willing to pay to get the medicine delivered. It's an imaginary situation, but please enter into the spirit of it. If you were told that it was going to cost you £10, would you say 'Yes, I'll pay that' or 'No, I'd rather put up with a couple of days with a bad sore throat'? Or would you be unsure?</p> <p style="text-align: right; font-size: small;"><i>Circle one answer</i></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="padding: 2px;">Yes</td> <td style="padding: 2px; text-align: center;">1</td> <td rowspan="3" style="padding: 2px;">GO TO B3 d</td> </tr> <tr> <td style="padding: 2px;">Unsure</td> <td style="padding: 2px; text-align: center;">2</td> </tr> <tr> <td style="padding: 2px;">No</td> <td style="padding: 2px; text-align: center;">3</td> <td style="padding: 2px;">GO TO B3 b</td> </tr> </table>	Yes	1	GO TO B3 d	Unsure	2	No	3	GO TO B3 b
Yes	1	GO TO B3 d							
Unsure	2								
No	3		GO TO B3 b						
B3 b	<p>READ OUT: Is there any amount less than £10 that you WOULD be prepared to pay to avoid a couple of days of this bad sore throat? If so, what is the most you would be prepared to pay?</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="padding: 2px;">Yes, the most would pay is: £</td> <td style="width: 100px; height: 20px;"></td> <td style="padding: 2px;">GO TO B4</td> </tr> <tr> <td style="padding: 2px;">No, I'm not prepared to pay anything</td> <td style="text-align: center; padding: 2px;">2</td> <td style="padding: 2px;">GO TO B3 c</td> </tr> </table>	Yes, the most would pay is: £		GO TO B4	No, I'm not prepared to pay anything	2	GO TO B3 c		
Yes, the most would pay is: £		GO TO B4							
No, I'm not prepared to pay anything	2	GO TO B3 c							
B3 c	<p>Please can you tell me why this cure would not be worth anything to you?</p> <p>INTERVIEWER WRITE IN REASONS BELOW:</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="width: 500px; height: 40px;"></td> <td style="padding: 2px;">GO TO B4</td> </tr> </table>		GO TO B4						
	GO TO B4								
B3 d	<p>READ OUT: Suppose instead that you were told that it was going to cost you £100? What would you say now? 'Yes, I'll pay that' or 'No, I'd rather put up with a couple of days with a sore throat' or would you be unsure?</p> <p style="text-align: right; font-size: small;"><i>Circle one answer</i></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="padding: 2px;">Yes</td> <td style="padding: 2px; text-align: center;">1</td> <td rowspan="3" style="padding: 2px;">GO TO B3 e</td> </tr> <tr> <td style="padding: 2px;">Unsure</td> <td style="padding: 2px; text-align: center;">2</td> </tr> <tr> <td style="padding: 2px;">No</td> <td style="padding: 2px; text-align: center;">3</td> </tr> </table>	Yes	1	GO TO B3 e	Unsure	2	No	3	
Yes	1	GO TO B3 e							
Unsure	2								
No	3								
B3 e	<p>So what is the MOST you would be prepared to pay to avoid those 2 days with a painful sore throat?</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="padding: 2px;">Most would pay is: £</td> <td style="width: 100px; height: 20px;"></td> <td style="padding: 2px;">GO TO B4</td> </tr> </table>	Most would pay is: £		GO TO B4					
Most would pay is: £		GO TO B4							

For office use only

1051

1052-1054

1055-1070

1071

1072-1078

PART B: (cont...)

Spare

B4 **READ OUT:** OK, now I've explained the general idea, I am going to ask you what value you would put on avoiding some of the illnesses you read about a few minutes ago.

GIVE SHOWCARD A TO THE RESPONDENT

B4 a Please read this card

AFTER RESPONDENT HAS READ DESCRIPTION, READ OUT:

Now imagine a test shows that you will shortly suffer what is described on **showcard A**. However, suppose that you could pay for a simple, safe and painless cure that would avoid everything on that card.

When you are thinking about what it would be worth to you to avoid all that, please ignore any loss of income that might be caused by the illness – suppose that your income is unaffected and just focus on how the illness would affect your health.

What I mean by your health are the things you considered in the first question such as how your usual work and leisure activities are affected, as well as any pain or discomfort or depression the illness might involve.

So thinking just about how your **health** would be affected, would avoiding everything on the card be worth at least SOMETHING to you, even just a few pence?

Circle one answer

Yes	1	GO TO B4 c
No	2	GO TO B4 b

1081

B4 b Could you say why this cure would **not** be worth **anything** at all to you?
(WRITE IN REASONS)

1082-
1097**GO TO B5**

B4 c **PICK UP MONEY CARDS & GIVE PAYMENT SHEET TO RESPONDENT**

READ OUT: I have a set of cards with different sums of money printed on them, and I'm going to ask you to sort them into piles on this sheet.

In this section (**INDICATE ON PAYMENT SHEET**) I'd like you to put all the amounts you **definitely would** be prepared to pay to avoid everything on that card. In this section (**INDICATE**) I'd like you to put all the amounts you definitely would **not** be prepared to pay. And if there are any amounts you are unsure about, we'll put them in this middle section and come back to them later.

As I show you each amount, please tell me which section you want me to put it into. Some of the cards involve quite large sums of money. In those cases, where you might not have the amount readily available, please think about whether avoiding the illness is valuable enough to you to be worth taking out a loan or extending your mortgage or something similar.

When you are thinking about this, please bear in mind what you **really** think you would or would not be prepared to pay, given your actual income and savings.

B4 d

SHUFFLE THE MONEY CARDS

READ OUT: I'm shuffling the cards so that they are in no particular order.

TAKE FIRST CARD FROM THE TOP OF THE PILE AND READ OUT:

Suppose it would cost you (*amount on card*) to get an instant cure and avoid everything on that card. Would you definitely **pay** this amount, definitely **not** pay it, or are you unsure?

PUT CARD IN APPROPRIATE SECTION

B4 e

TAKE NEXT CARD FROM THE TOP OF THE PILE AND READ OUT:

Suppose it would cost you (*amount on card*) to get an instant cure and avoid everything on that card. Would you definitely **pay** this amount, definitely **not** pay it, or feel unsure?

PUT CARD IN APPROPRIATE SECTION.

ONCE RESPONDENT UNDERSTANDS THE PROCESS, HAND THE PILE OF CARDS TO THEM AND READ OUT:

Now please work your way through the rest of the money amounts, sorting them according to how you feel.

B4 f

WAIT UNTIL RESPONDENT HAS SORTED ALL CARDS. IF ANY CARDS IN "UNSURE" PILE ASK IF RESPONDENT WANTS TO RE-ALLOCATE.

WHEN CARDS IN APPROPRIATE PILES RECORD:

Highest amount in "would pay" pile	£		GO TO B4 g
Lowest amount in "would not pay" pile	£		

Spare
1098-
1099

1100-
1106
1107-
1113

B4 g

We only gave you 16 amounts of money on those cards, so there were lots of gaps. The card you identified as the highest amount you definitely **WOULD** pay was: (*read out first amount from B4 f*). Would you actually be prepared to pay more than that?

Circle one answer

Yes	1	GO TO B4 h
No	2	GO TO B5

1114

B4 h

What is the **MOST** you would be prepared to pay to avoid everything that is on card **A**?

Most would pay is:	£		GO TO B5
--------------------	---	--	-----------------

1115-
1121

B5	Now let's do the same thing for this illness. GIVE SHOWCARD B TO THE RESPONDENT								
B5 a	Please read this card. AFTER RESPONDENT HAS READ DESCRIPTION, READ OUT: In this case, the illness would go on for 12 months rather than the 3 months in the previous case. Again, suppose you could get a simple, painless and safe cure that will prevent everything on card B . As before, please ignore completely any loss of income that might be caused by this illness and just focus on the effect on your health . Would avoiding everything on the card be worth at least SOMETHING to you, even just a few pence? <p style="text-align: right;"><i>Circle one answer</i></p> <table border="1" style="margin-left: auto;"> <tr> <td>Yes</td> <td style="text-align: center;">1</td> <td>GO TO B5 c</td> </tr> <tr> <td>No</td> <td style="text-align: center;">2</td> <td>GO TO B5 b</td> </tr> </table>	Yes	1	GO TO B5 c	No	2	GO TO B5 b	1122	
Yes	1	GO TO B5 c							
No	2	GO TO B5 b							
B5 b	Could you say why this cure would not be worth anything at all to you? (WRITE IN REASONS) <table border="1" style="width: 100%; height: 40px;"> <tr> <td style="width: 80%;"></td> <td style="text-align: right;">GO TO B6</td> </tr> </table>		GO TO B6	1123-1138					
	GO TO B6								
B5 c	SHUFFLE THE MONEY CARDS READ OUT: I'm going to use the same cards, but I'm shuffling them so that they are in no particular order. HAND SHUFFLED CARDS TO RESPONDENT As before, please work your way through these cards, sorting them according to whether you definitely would or definitely would not pay, and put any you are not sure about in the middle and come back to them later.								
B5 d	WAIT UNTIL RESPONDENT HAS SORTED ALL CARDS. IF ANY CARDS IN "UNSURE" PILE ASK IF RESPONDENT WANTS TO RE-ALLOCATE. WHEN CARDS IN APPROPRIATE PILES RECORD: <table border="1" style="width: 100%;"> <tr> <td style="width: 40%;">Highest amount in "would pay" pile</td> <td style="width: 10%;">£</td> <td style="width: 30%;"></td> <td rowspan="2" style="width: 20%; text-align: right;">GO TO B5 e</td> </tr> <tr> <td>Lowest amount in "would not pay" pile</td> <td>£</td> <td></td> </tr> </table>	Highest amount in "would pay" pile	£		GO TO B5 e	Lowest amount in "would not pay" pile	£		1139-1145 1146-1152
Highest amount in "would pay" pile	£		GO TO B5 e						
Lowest amount in "would not pay" pile	£								
B5 e	The card you identified as the highest amount you definitely WOULD pay was: (read out first amount from B5 d). Would you actually be prepared to pay more than that? <p style="text-align: right;"><i>Circle one answer</i></p> <table border="1" style="margin-left: auto;"> <tr> <td>Yes</td> <td style="text-align: center;">1</td> <td>GO TO B5 f</td> </tr> <tr> <td>No</td> <td style="text-align: center;">2</td> <td>GO TO B6</td> </tr> </table>	Yes	1	GO TO B5 f	No	2	GO TO B6	1153	
Yes	1	GO TO B5 f							
No	2	GO TO B6							
B5 f	What is the MOST you would be prepared to pay to avoid everything on card B ? <table border="1" style="width: 100%;"> <tr> <td style="width: 40%;">Most would pay is:</td> <td style="width: 10%;">£</td> <td style="width: 30%;"></td> <td style="width: 20%; text-align: right;">GO TO B6</td> </tr> </table>	Most would pay is:	£		GO TO B6	1154-1160			
Most would pay is:	£		GO TO B6						

For office use only

PART B: SECTION 2									
B6	<p>In the two questions you have just answered, I asked you to imagine that you were certain to get ill unless you were treated. Now suppose you are not certain to get ill, but that a test shows there is a chance you will become ill. This time, suppose a test suggests that there is a 5% chance that an illness will develop during the next few days. Of course, this means there is also a 95% chance that you will not be affected at all, even if you take no further action.</p> <p>Let us start by looking at the illness that lasts 3 months shown on showcard A again.</p> <p>GIVE CARD A TO THE RESPONDENT AND ALLOW TIME TO READ IT</p>								
B6 a	<p>As before, please ignore any loss of income that might be caused if you get ill, and just focus on the effect of the illness on your health.</p> <p>There is a 95% chance you will NOT get ill, but there is a 5% chance you will suffer the things shown on card A. Suppose there is a simple, painless and safe treatment that will eliminate that 5% chance.</p> <p>Would eliminating that 5% chance be worth at least SOMETHING to you, even just a few pence?</p>	1161							
<p style="margin: 0;"><i>Circle one answer</i></p> <table border="1" style="display: inline-table; border-collapse: collapse;"> <tr> <td style="padding: 2px 5px;">Yes</td> <td style="padding: 2px 5px; text-align: center;">1</td> <td style="padding: 2px 5px;">GO TO B6 c</td> </tr> <tr> <td style="padding: 2px 5px;">No</td> <td style="padding: 2px 5px; text-align: center;">2</td> <td style="padding: 2px 5px;">GO TO B6 b</td> </tr> </table>		Yes	1	GO TO B6 c	No	2	GO TO B6 b		
Yes	1	GO TO B6 c							
No	2	GO TO B6 b							
B6 b	<p>Could you say why this would not be worth anything to you? (WRITE IN REASONS)</p> <div style="border: 1px solid black; height: 40px; width: 100%;"></div>	1162-1177							
GO TO B7									
B6 c	<p>SHUFFLE THE MONEY CARDS AND HAND TO RESPONDENT</p> <p>READ OUT: As before, please sort the cards.</p>								
B6 d	<p>WHEN SORTED, ASK IF RESPONDENT WANTS TO RE-ALLOCATE ANY CARDS IN "UNSURE" PILE.</p> <p>THEN RECORD:</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 5px;">Highest amount in "would pay" pile</td> <td style="padding: 5px;">£</td> <td style="padding: 5px;"></td> <td rowspan="2" style="padding: 5px; text-align: center;">GO TO B6 e</td> </tr> <tr> <td style="padding: 5px;">Lowest amount in "would not pay" pile</td> <td style="padding: 5px;">£</td> <td style="padding: 5px;"></td> </tr> </table>	Highest amount in "would pay" pile	£		GO TO B6 e	Lowest amount in "would not pay" pile	£		1178-1184 1185-1191
Highest amount in "would pay" pile	£		GO TO B6 e						
Lowest amount in "would not pay" pile	£								
B6 e	<p>The highest card in this pile is: (read out first amount from B6 d). Would you actually be prepared to pay more than that to eliminate the 5% chance of that 3-month illness?</p>	1192							
<p style="margin: 0;"><i>Circle one answer</i></p> <table border="1" style="display: inline-table; border-collapse: collapse;"> <tr> <td style="padding: 2px 5px;">Yes</td> <td style="padding: 2px 5px; text-align: center;">1</td> <td style="padding: 2px 5px;">GO TO B6 f</td> </tr> <tr> <td style="padding: 2px 5px;">No</td> <td style="padding: 2px 5px; text-align: center;">2</td> <td style="padding: 2px 5px;">GO TO B7</td> </tr> </table>		Yes	1	GO TO B6 f	No	2	GO TO B7		
Yes	1	GO TO B6 f							
No	2	GO TO B7							
B6 f	<p>What is the MOST you would be prepared to pay to eliminate a 5% chance of that illness?</p>	1193-1199							
<p>Most would pay is: £</p> <table border="1" style="display: inline-table; border-collapse: collapse;"> <tr> <td style="padding: 5px;"></td> <td style="padding: 5px;"></td> <td style="padding: 5px;">GO TO B7</td> </tr> </table>				GO TO B7					
		GO TO B7							

B7	Now I'd like you to do the same thing again for the illness that goes on for 12 months.							
GIVE SHOWCARD B TO THE RESPONDENT AND ALLOW TIME TO READ IT								
B7 a	Suppose there is a 5% chance of this illness. Ignoring any possible loss of income, and focusing only on the effect of the illness on your health, suppose there is a simple, painless and safe treatment that will eliminate that 5% chance.							
Would eliminating that 5% chance be worth at least SOMETHING to you?								
<i>Circle one answer</i>								
<table border="1"> <tr> <td>Yes</td> <td>1</td> <td>GO TO B7 c</td> </tr> <tr> <td>No</td> <td>2</td> <td>GO TO B7 b</td> </tr> </table>			Yes	1	GO TO B7 c	No	2	GO TO B7 b
Yes	1	GO TO B7 c						
No	2	GO TO B7 b						
B7 b	Could you say why it would not be worth anything to you? (WRITE IN REASONS)	1201-1216						
		GO TO B8						
B7 c	SHUFFLE THE MONEY CARDS AND HAND TO RESPONDENT							
READ OUT: As before, please sort the cards.								
B7 d	WHEN SORTED, ASK IF RESPONDENT WANTS TO RE-ALLOCATE ANY CARDS IN "UNSURE" PILE.							
THEN RECORD:								
Highest amount in "would pay" pile		£						
Lowest amount in "would not pay" pile		£						
		GO TO B7 e						
B7 e	The highest card in the WOULD pay pile is: (read out first amount from B7 d). Would you actually be prepared to pay more than that to eliminate the 5% chance?	1217-1223						
<i>Circle one answer</i>								
<table border="1"> <tr> <td>Yes</td> <td>1</td> <td>GO TO B7 f</td> </tr> <tr> <td>No</td> <td>2</td> <td>GO TO B8</td> </tr> </table>			Yes	1	GO TO B7 f	No	2	GO TO B8
Yes	1	GO TO B7 f						
No	2	GO TO B8						
B7 f	What is the MOST you would be prepared to pay?	1231						
Most would pay is:		£						
		GO TO B8						

B8 Now suppose that if you get ill, it would involve episodes every couple of weeks for the rest of your life, as described on **showcard C**.

GIVE CARD C TO THE RESPONDENT AND ALLOW TIME TO READ IT

B8 a As before, there is a 95% chance you will not be affected, but there is a **5%** chance of developing **this** illness. Ignoring any loss of income and focusing only on the effect of the illness on your health:

Would eliminating that 5% chance be worth at least SOMETHING to you?

Circle one answer

Yes	1	GO TO B8 c
No	2	GO TO B8 b

B8 b Could you say why this would **not** be worth **anything** to you?
(WRITE IN REASONS)

	GO TO C1
--	-----------------

B8 c **SHUFFLE THE MONEY CARDS AND HAND TO RESPONDENT**

As before, please sort the cards.

B8 d **ASK IF RESPONDENT WANTS TO RE-ALLOCATE ANY “UNSURE”.**

WHEN CARDS IN APPROPRIATE PILES RECORD:

Highest amount in “would pay” pile	£		GO TO B8 e
Lowest amount in “would not pay” pile	£		

B8 e Would you actually be prepared to pay more than *(read out first amount from B8 d)*?

Circle one answer

Yes	1	GO TO B8 f
No	2	GO TO C1

B8 f What is the **MOST** you would be prepared to pay to eliminate that 5% chance?

Most would pay is:

£		GO TO C1
---	--	-----------------

PART C: (Interviewers to administer)

C1 READ OUT: The next few questions ask you to think about treatments with different chances of success and failure.

Some treatments give a very predictable outcome. But there are other treatments that may be more uncertain: sometimes they work well and improve people's health, but sometimes they fail and actually make things worse.

Which treatment you choose will probably depend on the chances of success or failure.

**C2 GIVE SHOWCARD G TO RESPONDENT.
POINT TO RELEVANT PARTS ON THE SHOWCARDS.**

READ OUT: The treatment on the LEFT (*indicate*) is totally predictable: if you go for that one, you will recover from a bad sore throat after 3 days for **CERTAIN**. The outcome of the treatment on the RIGHT is **UNCERTAIN** (*indicate*).

Suppose there is a 90% chance that the uncertain treatment succeeds and gets rid of the sore throat within 24 hours, but there is also a 10% chance that it fails and make things worse, so that the sore throat lasts for a full week.

Which treatment would you pick? 3 days with a bad sore throat for sure? Or a 90% chance of getting better within 24 hours but a 10% chance of having a bad sore throat for a full week?

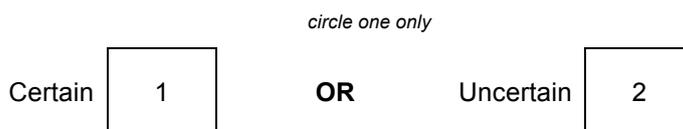
Certain	1	GO TO C3
Uncertain	2	

1278

C3 | GIVE SHOWCARD L TO RESPONDENT

READ OUT: Suppose that the predictable treatment means that you will **certainly** experience the effects described on the left. That is, you will have episodes every couple of weeks for 3 months. Alternatively, the treatment on the right gives you a **90%** chance of avoiding the illness altogether. BUT, there is also a **10%** chance that this treatment might leave you worse off, so that you will experience episodes every couple of weeks for 12 months rather than for 3 months.

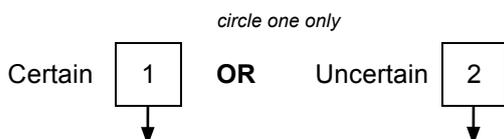
If those were the chances, which would you choose?
The certainty of the 3-month illness, or a 10% chance of the 12-month illness?



C4:

READ OUT: Now suppose instead that the **uncertain** treatment has a **99%** chance of success and a **1%** chance of failure. If **those** were the chances, which would you choose: the certainty of a 3 month illness or a 1% chance of a 12 month illness?

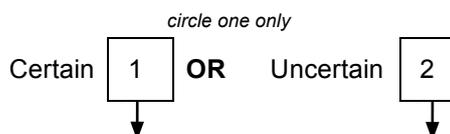
1280



C5:

READ OUT: Now suppose instead that the **uncertain** treatment had a **50%** chance of success and a **50%** chance of failure. If **those** were the chances, which would you choose: the certainty of a 3 month illness or a 50% chance of a 12 month illness?

1279



1281

C6:
What if the uncertain treatment had a **1 in 1000** chance of failure? If **that** was the chance of failure, which would you choose?

Certain	1	GO TO C10
Uncertain	2	GO TO D1

1282

C7:
What if the uncertain treatment had a **5%** chance of failure? If **that** was the chance of failure, which would you choose?

Certain	1	GO TO C11
Uncertain	2	GO TO C12

1283

C8:
What if the chances of failure for the uncertain treatment were:

READ OUT: STOP WHEN U PICKED

Circle C or U

- 40% **C** ₁ **U** ₂ 1284
- 30% **C** ₁ **U** ₂ 1285
- 20% **C** ₁ **U** ₂ 1286

GO TO D1

C9:
What if the chances of failure for the uncertain treatment were:

READ OUT: STOP WHEN C PICKED

Circle C or U

- 60% **C** ₁ **U** ₂ 1287
- 75% **C** ₁ **U** ₂ 1288
- 90% **C** ₁ **U** ₂ 1289
- 95% **C** ₁ **U** ₂ 1290
- 99% **C** ₁ **U** ₂ 1291

GO TO D1

- C10 READ OUT:** You said you would not take the uncertain treatment if the chance of failure was 1 in a 1000. How small would the chance of failure need to be for you to consider taking that treatment rather than accepting the certainty of the 3-month illness?

Would you be willing to take the uncertain treatment if the chance of failure was...**READ OUT...**
(STOP READING OUT IF RESPONDENT ANSWERS YES).

	Yes	No	
...one in ten thousand,	1	2	GO TO D1
one in a hundred thousand,	1	2	
one in a million?	1	2	

1292

1293

1294

- C11 READ OUT:** You said you WOULD take the uncertain treatment if the chance of failure was 1% but you would NOT take it when the chance of failure was 5%. Could you say what chances of failure between 1% and 5% you would accept? Would you take the uncertain treatment if the chance of failure was...**READ OUT...**
(STOP READING OUT IF RESPONDENT ANSWERS YES).

	Yes	No	
...4%,	1	2	GO TO D1
3%,	1	2	
or 2%?	1	2	

1295

1296

1297

- C12 READ OUT:** You said you WOULD take the uncertain treatment if the chance of failure was 5% but you would NOT take it when the chance of failure was 10%. Could you say what chances of failure between 5% and 10% you would accept? Would you take the uncertain treatment if the chance of failure was...**READ OUT...**
(STOP READING OUT IF RESPONDENT ANSWERS YES).

	Yes	No	
...9%,	1	2	GO TO D1
8%,	1	2	
7%	1	2	
Or 6%?	1	2	

1298

1299

1300

1301

D8	<p>READ OUT: You said you would not take the UNCERTAIN treatment if the chance of failure was 1 in a thousand. How small would the chance of failure need to be for you to consider taking that treatment rather than accepting the certainty of the 12-month illness?</p>												
	<p>Would you be willing to take the uncertain treatment if the chance of failure was...READ OUT... (STOP READING OUT IF RESPONDENT ANSWERS YES).</p>												
	<p>...one in ten thousand, one in hundred thousand, or one in a million?</p>	<table border="1"> <thead> <tr> <th>Yes</th> <th>No</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> </tbody> </table>	Yes	No	1	2	1	2	1	2	<p>1315 1316 1317</p> <p>GO TO E1</p>		
Yes	No												
1	2												
1	2												
1	2												
D9	<p>READ OUT: You said you WOULD take the uncertain treatment if the chance of failure was 1% but you would NOT take it when the chance of failure was 5%. Could you say what chances of failure between 1% and 5% you would accept? Would you take the uncertain treatment if the chance of failure was...READ OUT... (STOP READING OUT IF RESPONDENT ANSWERS YES).</p>												
	<p>...4%, 3%, or, 2%?</p>	<table border="1"> <thead> <tr> <th>Yes</th> <th>No</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> </tbody> </table>	Yes	No	1	2	1	2	1	2	<p>1318 1319 1320</p> <p>GO TO E1</p>		
Yes	No												
1	2												
1	2												
1	2												
D10	<p>READ OUT: You said you WOULD take the uncertain treatment if the chance of failure was 5% but you would NOT take it when the chance of failure was 10%. Could you say what chances of failure between 5% and 10% you would accept? Would you take the uncertain treatment if the chance of failure was...READ OUT... (STOP READING OUT IF RESPONDENT ANSWERS YES).</p>												
	<p>...9%, 8%, 7%, or, 6%?</p>	<table border="1"> <thead> <tr> <th>Yes</th> <th>No</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> </tbody> </table>	Yes	No	1	2	1	2	1	2	1	2	<p>1321 1322 1323 1324</p> <p>GO TO E1</p>
Yes	No												
1	2												
1	2												
1	2												
1	2												

Spare 1325 - 1330

- E8** **READ OUT:** You said you would not take the uncertain treatment if the chance of failure was 1 in a thousand. How small would the chance of failure need to be for you to consider taking that treatment rather than accepting the certainty of the illness for the rest of your life?

Would you be willing to take the treatment on the right if the chance of failure was...**READ OUT...**

(STOP READING OUT IF RESPONDENT ANSWERS YES).

	Yes	No		
...one in ten thousand,	1	2	GO TO F1	1344
one in hundred thousand,	1	2		1345
one in a million?	1	2		1346

- E9** **READ OUT:** You said you WOULD take the uncertain treatment if the chance of failure was 1% but you would NOT take it when the chance of failure was 5%. Could you say what chances of failure between 1% and 5% you would accept? Would you take the uncertain treatment if the chance of failure was...**READ OUT...**

(STOP READING OUT IF RESPONDENT ANSWERS YES).

	Yes	No		
...4%,	1	2	GO TO F1	1347
3%,	1	2		1348
or, 2%?	1	2		1349

- E10** **READ OUT:** You said you WOULD take the uncertain treatment if the chance of failure was 5% but you would NOT take it when the chance of failure was 10%. Could you say what chances of failure between 5% and 10% you would accept? Would you take the uncertain treatment if the chance of failure was...**READ OUT...**

(STOP READING OUT IF RESPONDENT ANSWERS YES).

	Yes	No		
...9%,	1	2	GO TO F1	1350
8%,	1	2		1351
7%	1	2		1352
or, 6%?	1	2		1353

F8	<p>READ OUT: You said you would not take the uncertain treatment if the chance of failure was 1 in a thousand. How small would the risk of failure need to be for you to consider taking that treatment rather than accepting the certainty of a year of normal health followed by having the illness for the rest of your life?</p> <p>Would you be willing to take the treatment on the right if the risk of failure was...READ OUT... (STOP READING OUT IF RESPONDENT ANSWERS YES).</p>	<table border="1"> <thead> <tr> <th>Yes</th> <th>No</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> </tbody> </table>	Yes	No	1	2	1	2	1	2	<p>...one in ten thousand, one in hundred thousand, one in a million?</p> <p>GO TO G1</p>	<p>1367 1368 1369</p>		
Yes	No													
1	2													
1	2													
1	2													
F9	<p>READ OUT: You said you WOULD take the uncertain treatment if the chance of failure was 1% but you would NOT take it when the chance of failure was 5%. Could you say what chances of failure between 1% and 5% you would accept? Would you take the uncertain treatment if the chance of failure was...READ OUT... (STOP READING OUT IF RESPONDENT ANSWERS YES).</p>	<table border="1"> <thead> <tr> <th>Yes</th> <th>No</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> </tbody> </table>	Yes	No	1	2	1	2	1	2	<p>...4%, 3%, or, 2%?</p> <p>GO TO G1</p>	<p>1370 1371 1372</p>		
Yes	No													
1	2													
1	2													
1	2													
F10	<p>READ OUT: You said you WOULD take the uncertain treatment if the chance of failure was 5% but you would NOT take it when the chance of failure was 10%. Could you say what chances of failure between 5% and 10% you would accept? Would you take the uncertain treatment if the chance of failure was...READ OUT... (STOP READING OUT IF RESPONDENT ANSWERS YES).</p>	<table border="1"> <thead> <tr> <th>Yes</th> <th>No</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> <tr> <td>1</td> <td>2</td> </tr> </tbody> </table>	Yes	No	1	2	1	2	1	2	1	2	<p>...9%, 8%, 7% or, 6%?</p> <p>GO TO G1</p>	<p>1373 1374 1375 1376</p>
Yes	No													
1	2													
1	2													
1	2													
1	2													

Spare 1377-1446

For office use only

PART G: SECTION 1			
G1	How many people (including adults and children) live in your household?		
	One only	1	GO TO G11
	2 or more	2	GO TO G2
G2	MULTI-OCCUPANCY HOUSEHOLDS ONLY		
	We are interested in the things you were thinking about when you were answering the 'money questions'. These were the ones where we asked you to sort the money cards into piles to show how much you would be willing to pay for various treatments. INTERVIEWER: IF NECESSARY, SHOW MONEY CARDS TO REMIND RESPONDENT		
	When you answered these questions, were you thinking about income and savings?		
	Yes	1	GO TO G3
	No	2	GO TO G4
G3	Were you thinking about...READ OUT...		
	... The whole of your household income,	1	GO TO G4
	...your own personal income	2	
	Or, a fraction of your household income?	3	
G4	And, when answering these questions were you taking into account any financial commitments that you may have?		
	Yes	1	GO TO G5
	No	2	GO TO G7

1447

1448

1449

1450



G5	Which of the following, if any, were you thinking about when answering the money sort questions...READ OUT...					
		Yes	No	N/A		
	Mortgage or rent payments	1	2	3	GO TO G7	1451
	Debts e.g. credit card debts, loans	1	2	3		1452
	Child support or childcare costs	1	2	3		1453
	Cost of children's education e.g. university fees	1	2	3		1454
	Regular bills e.g. gas, telephone etc	1	2	3		1455
Or, other financial commitments?	1	2	3	If yes, GO TO G6		1456
G6	What other financial commitments were you thinking about? INTERVIEWER WRITE IN BELOW				1457-1472	
G7	When answering the money sort questions, were you thinking about the wellbeing of other people who are dependent on you, who may live either in your household or elsewhere?					
	Yes	1	GO TO G8		1473	
	No	2	GO TO G9			
G8	Were you thinking about...READ OUT...					
		Yes	No	N/A		
	Your spouse or partner,	1	2	3	GO TO G9	1474
	Your children	1	2	3		1475
	Other dependants currently living in your household	1	2	3		1476
Or, other dependants who live elsewhere	1	2	3	1477		

G9

Looking at the things listed on Showcard P, which was the most important factor when answering how much you would be willing to pay for a treatment?

1478-1479

Wellbeing of:	Your personal or household income/ savings	1	GO TO G19	
	Spouse or partner	2		
	Your children	3		
	Other dependants currently living in your household	4		
	Or, other dependants who live elsewhere	5		
	Financial commitments:	Mortgage or rent payments		6
		Debts e.g. credit card debts, loans		7
		Childcare costs		8
		Cost of children's education e.g. university fees		9
		Regular bills e.g. gas, telephone etc		10
		Other financial commitments		11
		Or something else		12

G10

What was that? INTERVIEWER WRITE IN BELOW

1480-1495

GO TO G19

G11

SINGLE OCCUPANCY HOUSEHOLDS ONLY:

We are interested in the things you were thinking about when you were answering the 'money questions'. These were the ones where we asked you to sort the money cards into piles to show how much you would be willing to pay for various treatments.

INTERVIEWER: IF NECESSARY, SHOW MONEY CARDS TO REMIND RESPONDENT

When you answered these questions, were you thinking about income and savings?

Yes	1	GO TO G12
No	2	

1496

G12	And, when answering these questions were you taking into account any financial commitments that you may have?	Yes	1	GO TO G13	1497		
		No	2	GO TO G14			
G13	Which of the following, if any, were you thinking about when answering the money sort questions...READ OUT...	Yes	No	N/A	GO TO G15 If yes, GO TO G14	1498 1499 1500 1501 1502 1503	
		Mortgage or rent payments	1	2			3
		Debts e.g. credit card debts, loans	1	2			3
		Childcare costs	1	2			3
		Cost of children's education e.g. university fees	1	2			3
		Regular bills e.g. gas, telephone etc	1	2			3
		Or, other financial commitments?	1	2			3
G14	What other financial commitments were you thinking about? INTERVIEWER WRITE IN BELOW				1504-1519		
G15	When answering the money sort questions, were you thinking about the wellbeing of other people who are dependent on you?	Yes	1	GO TO G16	1520		
		No	2	GO TO G17			
G16	Were you thinking about...READ OUT...	Yes	No	N/A	GO TO G17	1521 1522 1523	
		Your spouse or partner,	1	2			3
		Your children	1	2			3
		Or, other dependants who live elsewhere	1	2			3

G17

Looking at the things listed on Showcard P, which was the most important factor when answering how much you would be willing to pay for a treatment?

Wellbeing of:	Your personal or household income/ savings	1	GO TO G19	
	Spouse or partner	2		
	Your children	3		
	Or, other dependants who live elsewhere	5		
	Financial commitments:	Mortgage or rent payments		6
	Debts e.g. credit card debts, loans	7		
	Childcare costs	8		
	Cost of children's education e.g. university fees	9		
	Regular bills e.g. gas, telephone etc	10		
	Other financial commitments	11		
	Or something else	12		GO TO G18

1524-
1525**G18**

What was that? INTERVIEWER WRITE IN BELOW

1526-
1541

PART G: SECTION 2		
G19	<p>READ OUT</p> <p>Now I'd like to ask some questions about you which are needed to help interpret the answers you have just given me.</p>	
G20	<p>INTERVIEWER CODE: THE RESPONDENT SEX IS:</p> <p style="text-align: right;">Male <input type="text" value="1"/></p> <p style="text-align: right;">Female <input type="text" value="2"/></p>	1542
G21	<p>What was your age last birthday?</p> <p style="text-align: right;">ENTER AGE: <input type="text"/> <input type="text"/></p>	1543-1544
G22	<p>How is your health in general, would you say it is...READ OUT...</p> <p style="text-align: right;"><i>circle one code only</i></p> <p style="text-align: right;">...Very good, <input type="text" value="1"/></p> <p style="text-align: right;">Good, <input type="text" value="2"/></p> <p style="text-align: right;">Fair, <input type="text" value="3"/></p> <p style="text-align: right;">Bad, <input type="text" value="4"/></p> <p style="text-align: right;">Or, very bad? <input type="text" value="5"/></p>	1545
G23	<p>Do you have any longstanding illnesses. By longstanding I mean anything that has troubled you over a period of time or that is likely to affect you over a period of time?</p> <p style="text-align: right;"><i>Circle one code only</i></p> <p style="text-align: right;">Yes <input type="text" value="1"/> GO TO G24</p> <p style="text-align: right;">No <input type="text" value="2"/> GO TO G25</p>	1546
G24	<p>Does this longstanding illness limit your activities in any way?</p> <p style="text-align: right;"><i>Circle one code only</i></p> <p style="text-align: right;">Yes <input type="text" value="1"/> GO TO G25</p> <p style="text-align: right;">No <input type="text" value="2"/> GO TO G25</p>	1547

G25 **SHOWCARD Q.** Please tell me to which of these ethnic groups you consider that you belong?

1548-1549

Circle one code only

White	1	GO TO G26
White and Black Caribbean	2	
White and Black African	3	
White and Asian	4	
Any other mixed background	5	
Indian	6	
Pakistani	7	
Bangladeshi	8	
Any other Asian Background	9	
Caribbean	10	
African	11	
Any other black background	12	
Chinese	13	
other	14	

G26 Looking at **Showcard R**, do you have any of the qualifications listed on the card?

1550

Yes	1	GO TO G27
No	2	GO TO G28

G27 Which ones?**INTERVIEWER CODE ALL THAT APPLY**

Degree/degree level qualification	1	GCSE GRADES D-G	16
Teaching qualification	2	CSE GRADE 1/SCE BANDS A-C/Standard Grade LEVEL 1-3	17
Nursing qualifications SRN, SCM, SEN, RGN, RM, RHV, Midwife	3	CSE GRADES 2-5/SCE Ordinary BANDS D-E	18
HNC/HND, BEC/TEC Higher, BTEC etc	4	CSE Ungraded	19
ONC/OND/BEC/TEC/BTEC not higher	5	SLC Lower	20
City and Guilds Full Technological Certificate	6	SUPE Lower or Ordinary	21
City and Guilds Advanced/Final Level	7	School Certificate or Matric	22
City and Guilds Craft/Ordinary Level	8	NVQ Level 5	23
A-levels/Higher School Certificate	9	NVQ Level 4	24
AS level	10	NVQ Level 3/Advanced level GNVQ	25
SLC/SCE/SUPE at Higher Grade	11	NVQ Level 2/Intermediate level GNVQ	26
O-level passes taken in 1975 or earlier	12	NVQ Level 1/Foundation level GNVQ	27
O-level passes taken after 1975: A-C	13	Recognised Trade Apprenticeship	28
O-level taken after 1975 GRADES D-E	14	Clerical or Commercial Qualification (eg typing/ book-keeping/commerce)	29
GCSE GRADES A-C	15		

1551-1580

G28 Looking at **Showcard S**, which of these descriptions applies to what you were doing yesterday?*Circle one code only*

In paid employment or self-employment (or away temporarily)	1	GO TO G30
Waiting to take up paid work already obtained	2	
Looking for paid work or a Government training scheme	3	GO TO G29
Going to school or college full-time (including on vacation)	4	
Doing unpaid work for a business that you or a relative owns	5	
On a Government scheme for employment training	6	
Intending to look for work but prevented by temporary sickness or injury (sick or injured for 28 days or less)	7	
Permanently unable to work because of long-term sickness/disability	8	
Retired from paid work	9	
Looking after the home or family	10	
Doing something else	11	

1581-1582

G29 Have you ever had a paid job, apart from casual or holiday work?

Yes	1	GO TO G30
No	2	GO TO G39

1583

G30	<p>ASK ABOUT PRESENT JOB IF HRP IS CURRENTLY IN WORK, ASK ABOUT LAST JOB IF CURRENTLY NOT IN WORK</p> <p>What is/was the name or title of your job (in the week ending last Sunday)?</p> <div style="border: 1px solid black; height: 100px; width: 90%; margin-top: 10px;"></div> <p>ENTER JOB TITLE:</p>	<p>1584- 1586</p>						
G31	<p>What does/did the firm/organisation you work(ed) for mainly make or do (at the place where you/they work(ed))?</p> <p>DESCRIBE FULLY – PROBE MANUFACTURING or PROCESSING or DISTRIBUTING ETC.AND MAIN GOODS PRODUCED, MATERIALS USED, WHOLESALE or RETAIL ETC.</p> <div style="border: 1px solid black; height: 150px; width: 90%; margin-top: 10px;"></div>	<p><i>Spare</i> 1587- 1588</p> <p>1589- 1591</p>						
G32	<p>What kind of work do/did you do most of the time? INTERVIEWER CHECK: What skills or qualifications are/ were needed for the job?</p> <div style="border: 1px solid black; height: 150px; width: 90%; margin-top: 10px;"></div>	<p>1592- 1593</p>						
G33	<p>Are/were you working as an employee or are/were you self-employed?</p> <p style="text-align: right;"><i>Circle one code only</i></p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 60%; text-align: right;">An employee</td> <td style="width: 5%; text-align: center; border: 1px solid black;">1</td> <td style="width: 35%; text-align: left;">GO TO G34</td> </tr> <tr> <td style="text-align: right;">Self employed</td> <td style="text-align: center; border: 1px solid black;">2</td> <td style="text-align: left;">GO TO G36</td> </tr> </table>	An employee	1	GO TO G34	Self employed	2	GO TO G36	<p><i>Spare</i> 1594- 1595</p> <p>1596</p>
An employee	1	GO TO G34						
Self employed	2	GO TO G36						

G34	Are you a...READ OUT...	<i>circle one code only</i>		GO TO G35	1597
		...a manager,	1		
		Foreman or supervisor,	2		
		Or, other employee	3		
G35	How many people work(ed) for your employer at the place where you work(ed)?	<i>Circle one code only</i>		GO TO G38	1598
		1-24	1		
		25-499	2		
		500+	3		
		Can't say	4		
G36	Are/were you working on your own or do/did you have employees?	<i>Circle one code only</i>		GO TO G37	1599
		On own/with partner(s) but no employees	1		
		With employees	2		
G37	How many people do/did you employ at the place where you/they work(ed)	<i>circle one code only</i>		GO TO G38	1600
		1 or 2	1		
		3-24	2		
		25-499	3		
		500+	4		
		Can't say	5		
G38	In your (main) job are/were you working full or part time	<i>circle one code only</i>		GO TO G39	1601
		Full time	1		
		Part time	2		
G39	SHOWCARD T	This card shows incomes in weekly, monthly and annual amounts. Thinking of the income of your household as a whole, which of the groups on this card represents the total income of the whole household before deductions for income tax, National Insurance, etc? Just tell me the number beside the row that applies to your household. ENTER NUMBER HERE: <input type="text"/> <input type="text"/>			1602-1603

G40	Looking at Showcard U , Which of those statements best reflects your household's situation as regards money? Just tell me the number beside the statement that most applies to you.	1604
	<i>Circle one code only</i>	
	We normally have enough money for anything we want	1
	We have enough money, so long as we plan our spending carefully	2
	We have enough money for basic things, but we can't afford anything unnecessary	3
	Sometimes it is hard for us to afford even the basic things we need	4
G41	Would you be willing for a member of the research team to contact you in the future? As before, everything you say would be treated in complete confidence.	1605
	<i>circle one code only</i>	
	Yes	1
	No	2
G42	A certain number of interviews on any survey are checked by a supervisor to make sure that people were satisfied with the way the interview was carried out. Can we contact you for this purpose?	1606
	<i>circle one code only</i>	
	Yes	1
	No	2
		Enter phone number of front on ARF
G43	INTERVIEWER RECORD THE FOLLOWING INFORMATION? Total number of people resident in household aged 16 and over Total number of people resident in household aged 0-15	1607-1608 1609-1610
G44	That is the end of the interview, thank you very much for your help.	

Interviewer to complete:

1) Record end time of interview (use 24 hour clock) :

<i>Hours</i>			<i>mins</i>	
1611-1612			1613-1614	

2) Record Date of Interview: __ / __ / 2007

1615-1616 1617-1618

3) Interviewer signature: _____

Appendix 7

Show cards for ranking

M

Initially you will have stomach pains, diarrhoea and vomiting for **3–4 days**, severe enough to interfere with **some** of your usual activities.

Things then improve BUT during the next **3 months** you will suffer an episode of stomach discomfort and sickness **every couple of weeks**, with each episode lasting **2 or 3 days**.

These episodes are not so severe, but may still interfere with **some** of your usual activities.

After 3 months you return to your current health with no further effects from this illness.

E

Initially you will have **severe** stomach pains, diarrhoea, vomiting and fever for **7 days**, severe enough to interfere with **most** of your usual activities.

Things then improve BUT during the next **12 months** you will suffer an episode of stomach discomfort and sickness **every couple of weeks**, with each episode lasting **2 or 3 days**.

These episodes are not so severe, but may still interfere with **some** of your usual activities.

After 12 months you return to your current health with no further effects from this illness.

T

Initially you will have **severe** stomach pains, diarrhoea, vomiting and fever for **7 days**, severe enough to interfere with **most** of your usual activities.

Things then improve BUT **for the rest of your life** you will suffer an episode of stomach discomfort and sickness **every couple of weeks**, with each episode lasting **2 or 3 days**.

These episodes are not so severe, but may still interfere with **some** of your usual activities.

V

You will have episodes of throbbing pain across the front of the head and you will feel sick and may occasionally BE sick. You will feel like you want to lie still in a darkened room.

During the next **3 months** you will suffer an episode of head pain and sickness **every couple of weeks**, with each episode lasting **between 8 hours and 2 days**.

These episodes will interfere with **many** of your usual activities.

After 3 months you return to your current health with no further effects from this illness.

A

You will have episodes of throbbing pain across the front of the head and you will feel sick and may occasionally BE sick. You will feel like you want to lie still in a darkened room.

During the next **12 months** you will suffer an episode of head pain and sickness **every couple of weeks**, with each episode lasting **between 8 hours and 2 days**.

These episodes will interfere with **many** of your usual activities.

After 12 months you return to your current health with no further effects from this illness.

S

Every couple of weeks for the rest of your life you will have episodes of throbbing pain across the front of the head and you will feel sick and may occasionally BE sick. You will feel like you want to lie still in a darkened room.

Each episode will last **between 8 hours and 2 days**, and will interfere with **many** of your usual activities.

Appendix 8

Standard gamble cards

C

CERTAIN

FOR SURE

You will have episodes of throbbing pain across the front of your head and you will feel sick and may occasionally BE sick. You will feel like you want to lie still in a darkened room.

During the next **3 months** you will suffer an episode of head pain and sickness **every couple of weeks**, with each episode lasting **between 8 hours and 2 days**.

These episodes will interfere with **many** of your usual activities.

After these 3 months you return to your current health with no further effects from this illness.

UNCERTAIN

Your current state of health with no effects of this illness.

You will have episodes of throbbing pain across the front of your head and you will feel sick and may occasionally BE sick. You will feel like you want to lie still in a darkened room.

During the next **12 months** you will suffer an episode of head pain and sickness **every couple of weeks**, with each episode lasting **between 8 hours and 2 days**.

These episodes will interfere with **many** of your usual activities.

After these 12 months you return to your current health with no further effects from this illness.

D

CERTAIN

FOR SURE

You will have episodes of throbbing pain across the front of your head and you will feel sick and may occasionally BE sick. You will feel like you want to lie still in a darkened room.

During the next **12 months** you will suffer an episode of head pain and sickness **every couple of weeks**, with each episode lasting **between 8 hours and 2 days**.

These episodes will interfere with **many** of your usual activities.

After these 12 months you return to your current health with no further effects from this illness.

UNCERTAIN

Your current state of health with no effects of this illness.

Every couple of weeks for the rest of your life you will have episodes of throbbing pain across the front of the head and you will feel sick and may occasionally BE sick. You will feel like you want to lie still in a darkened room.

Each episode lasting **between 8 hours and 2 days**, and will interfere with **many** of your usual activities.



E

CERTAIN

FOR SURE

Every couple of weeks for the rest of your life you will have episodes of throbbing pain across the front of the head and you will feel sick and may occasionally BE sick. You will feel like you want to lie still in a darkened room.

Each episode lasting **between 8 hours and 2 days**, and will interfere with **many** of your usual activities.

UNCERTAIN

Your current state of health with no effects of this illness.

Immediate death

F

CERTAIN

FOR SURE

For the first 12 months after the treatment starts, you will continue in your current state of health. But after 12 months the treatment will stop being effective.

From then on, **every couple of weeks for the rest of your life** you will have episodes of throbbing pain across the front of the head and you will feel sick and may occasionally BE sick. You will feel like you want to lie still in a darkened room.

Each episode lasting **between 8 hours and 2 days**, and will interfere with **many** of your usual activities.

UNCERTAIN

Your current state of health with no effects of this illness.

Starting now, **every couple of weeks for the rest of your life** you will have episodes of throbbing pain across the front of the head and you will feel sick and may occasionally BE sick. You will feel like you want to lie still in a darkened room.

Each episode lasting **between 8 hours and 2 days**, and will interfere with **many** of your usual activities.



Appendix 9

Differences between discrete choice and matching

Discrete choices/group-level analysis	Characteristics of methods	Matching/individual-level analysis
PRESENTATION OF QUESTIONS TO RESPONDENTS		
Standard diagrams were used to present information to respondents. These diagrams were very similar whether discrete choices were presented or iterative matching questions. Respondents were able to assimilate information better in this way than in any other attempt we made to convey information in different formats. However the presentation of the questions varied in other ways:		
<p>1. Multiple attributes are varied in each choice set such that 'option A' and 'option B' represent groups of patients who are of different age groups and whose QoL differs. The size of the health gain (QALYs gained) is also allowed to vary. Because of this respondents are able to 'trade' the size of the health gain (in QALYs) against the age of the population in question, or their QoL.</p>	Varying and trading-off different attributes	<p>1. Typically only one attribute is varied at one time while everything else is held constant. This means that for any matching question 'option A' and 'option B' are of the same age but different QoL, or the same QoL and different ages. In all questions the health gain was of the same size in A as in B. Most presented gains of 4 QALYs (i.e. adjacent single cell comparisons in the QALY grid) with some 'whole row' and 'whole column' comparisons allowing for tests of additivity.</p>
<p>2. Choices are presented as one-off preferences for 'option A' or 'option B'. The points of indifference (relativities) between different scenarios are estimated indirectly using regression techniques.</p>	Points of indifference (deriving relativities)	<p>2. Respondents are presented with the same number of beneficiaries in 'option A' as in 'option B' and asked to choose. The number of beneficiaries in the most preferred option is then reduced systematically until an individual's point of indifference is reached.</p>
<p>3. The number of beneficiaries of the health gains is not stated. Respondents are told that there is the same number of beneficiaries in each group and the treatment costs the same in both.</p>	Number of beneficiaries of health gains	<p>3. Respondents are told that there are 100 people in each group initially. Depending on their choice of A or B, numbers are reduced in the preferred choice until they are indifferent between the two options. Number of beneficiaries is used as the numeraire.</p>

SURVEY DESIGN AND SELECTION OF QUESTIONS	
<p>Different design conventions are usually adopted for selection of discrete choice questions and selection of matching questions. The selection of discrete choice questions is perhaps more of an issue because of the estimation of relativities indirectly and so 'relativities' are not explicit at the individual level. The subject of this study and specifically the inter-relatedness of attributes presented serious challenges for the usual methods used to design discrete choice studies.</p>	
<p>4. A large number of attributes, when combined with levels on other attributes, resulted in impossible or implausible scenarios (for example 'age' and 'life-years gained' together giving an implausible age at death). As a result only 19% of total number of all possible scenarios could be used in the design.</p>	<p>4. The matching study essentially faces the same constraints in terms of the presentation of plausible choices – that is that questions must make sense (in terms of age, severity and any health gain).</p>
<p>5. Discrete choice studies typically make reference to methods of experimental design. Constraints imposed on the design of DCEs (such as outlined in 4) have an adverse effect on the flexibility of the analysis of the resultant data and the interpretation of the findings.</p>	<p>5. The design of matching questionnaires does not typically make reference to conventions of experimental design. Scenarios were paired therefore in such a way to cover the whole of the available space in a systematic way (i.e. presenting comparisons of all cells with adjacent cells).</p>
Interdependence of attributes and constraints on experimental design	
ANALYTIC ISSUES	
<p>The analyses of these two data sets are quite different. The advantages of the matching data are their 'transparency' at the individual level (a point of indifference is elicited for each individual respondent). In the SVQ study a 'QALY grid' approach has been taken to derive weights. The grid is divided into 20 cells each representing four QALYs. Weights are calculated for each of 19 cells relative to a reference cell. In discrete choice data relativities between two scenarios (described in terms of levels on five attributes) are inferred from aggregate data. This data set posed challenges for discrete choice analysis and the usual analytic models and techniques applied were not necessarily appropriate.</p>	
<p>6. Aggregation is not an issue for analysis of discrete choice data since an aggregate model is estimated (see point 7).</p>	<p>6. Results are dependent on the method of aggregation of ratios with different approaches to aggregation resulting in different weights.</p>
Aggregation	
<p>7. According to DCE convention, the experimental design dictates those models that can be estimated and so the results of this study are therefore accompanied by caveats. Results are also dependent on the assumptions underlying the functional form used in the econometric analysis.</p>	<p>7. The assumptions underlying the matching/QALY grid approach are largely also embedded in the QALY model itself.</p>
Assumptions	
<p>8. In principle, using discrete choice data, relative weights could be derived for all possible scenarios. General insights can, however, be drawn out of the data and reported qualitatively.</p>	<p>8. Results depend on the assumption that relative weights derived for a QALY gain of four QALYs can be generalised to QALY gains of all sizes. This is, of course, also an underlying assumption of the QALY model itself.</p>
Generalisation	

Health Technology Assessment reports published to date

Volume 1, 1997

No. 1

Home parenteral nutrition: a systematic review.

By Richards DM, Deeks JJ, Sheldon TA, Shaffer JL.

No. 2

Diagnosis, management and screening of early localised prostate cancer.

A review by Selley S, Donovan J, Faulkner A, Coast J, Gillatt D.

No. 3

The diagnosis, management, treatment and costs of prostate cancer in England and Wales.

A review by Chamberlain J, Melia J, Moss S, Brown J.

No. 4

Screening for fragile X syndrome.

A review by Murray J, Cuckle H, Taylor G, Hewison J.

No. 5

A review of near patient testing in primary care.

By Hobbs FDR, Delaney BC, Fitzmaurice DA, Wilson S, Hyde CJ, Thorpe GH, *et al.*

No. 6

Systematic review of outpatient services for chronic pain control.

By McQuay HJ, Moore RA, Eccleston C, Morley S, de C Williams AC.

No. 7

Neonatal screening for inborn errors of metabolism: cost, yield and outcome.

A review by Pollitt RJ, Green A, McCabe CJ, Booth A, Cooper NJ, Leonard JV, *et al.*

No. 8

Preschool vision screening.

A review by Snowdon SK, Stewart-Brown SL.

No. 9

Implications of socio-cultural contexts for the ethics of clinical trials.

A review by Ashcroft RE, Chadwick DW, Clark SRL, Edwards RHT, Frith L, Hutton JL.

No. 10

A critical review of the role of neonatal hearing screening in the detection of congenital hearing impairment.

By Davis A, Bamford J, Wilson I, Ramkalawan T, Forshaw M, Wright S.

No. 11

Newborn screening for inborn errors of metabolism: a systematic review.

By Seymour CA, Thomason MJ, Chalmers RA, Addison GM, Bain MD, Cockburn F, *et al.*

No. 12

Routine preoperative testing: a systematic review of the evidence.

By Munro J, Booth A, Nicholl J.

No. 13

Systematic review of the effectiveness of laxatives in the elderly.

By Petticrew M, Watt I, Sheldon T.

No. 14

When and how to assess fast-changing technologies: a comparative study of medical applications of four generic technologies.

A review by Mowatt G, Bower DJ, Brebner JA, Cairns JA, Grant AM, McKee L.

Volume 2, 1998

No. 1

Antenatal screening for Down's syndrome.

A review by Wald NJ, Kennard A, Hackshaw A, McGuire A.

No. 2

Screening for ovarian cancer: a systematic review.

By Bell R, Petticrew M, Luengo S, Sheldon TA.

No. 3

Consensus development methods, and their use in clinical guideline development.

A review by Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CFB, Askham J, *et al.*

No. 4

A cost-utility analysis of interferon beta for multiple sclerosis.

By Parkin D, McNamee P, Jacoby A, Miller P, Thomas S, Bates D.

No. 5

Effectiveness and efficiency of methods of dialysis therapy for end-stage renal disease: systematic reviews.

By MacLeod A, Grant A, Donaldson C, Khan I, Campbell M, Daly C, *et al.*

No. 6

Effectiveness of hip prostheses in primary total hip replacement: a critical review of evidence and an economic model.

By Faulkner A, Kennedy LG, Baxter K, Donovan J, Wilkinson M, Bevan G.

No. 7

Antimicrobial prophylaxis in colorectal surgery: a systematic review of randomised controlled trials.

By Song F, Glenny AM.

No. 8

Bone marrow and peripheral blood stem cell transplantation for malignancy.

A review by Johnson PWM, Simnett SJ, Sweetenham JW, Morgan GJ, Stewart LA.

No. 9

Screening for speech and language delay: a systematic review of the literature.

By Law J, Boyle J, Harris F, Harkness A, Nye C.

No. 10

Resource allocation for chronic stable angina: a systematic review of effectiveness, costs and cost-effectiveness of alternative interventions.

By Sculpher MJ, Petticrew M, Kelland JL, Elliott RA, Holdright DR, Buxton MJ.

No. 11

Detection, adherence and control of hypertension for the prevention of stroke: a systematic review.

By Ebrahim S.

No. 12

Postoperative analgesia and vomiting, with special reference to day-case surgery: a systematic review.

By McQuay HJ, Moore RA.

No. 13

Choosing between randomised and nonrandomised studies: a systematic review.

By Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C.

No. 14

Evaluating patient-based outcome measures for use in clinical trials.

A review by Fitzpatrick R, Davey C, Buxton MJ, Jones DR.

No. 15

Ethical issues in the design and conduct of randomised controlled trials.

A review by Edwards SJL, Lilford RJ, Braunholtz DA, Jackson JC, Hewison J, Thornton J.

No. 16

Qualitative research methods in health technology assessment: a review of the literature.

By Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P.

No. 17

The costs and benefits of paramedic skills in pre-hospital trauma care.

By Nicholl J, Hughes S, Dixon S, Turner J, Yates D.

No. 18

Systematic review of endoscopic ultrasound in gastro-oesophageal cancer.

By Harris KM, Kelly S, Berry E, Hutton J, Roderick P, Cullingworth J, *et al.*

No. 19

Systematic reviews of trials and other studies.

By Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F.

No. 20

Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses.

A review by Fitzpatrick R, Shortall E, Sculpher M, Murray D, Morris R, Lodge M, *et al.*

Volume 3, 1999

No. 1

Informed decision making: an annotated bibliography and systematic review.

By Bekker H, Thornton JG, Airey CM, Connelly JB, Hewison J, Robinson MB, *et al.*

No. 2

Handling uncertainty when performing economic evaluation of healthcare interventions.

A review by Briggs AH, Gray AM.

No. 3

The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review.

By Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H.

No. 4

A randomised controlled trial of different approaches to universal antenatal HIV testing: uptake and acceptability. Annex: Antenatal HIV testing – assessment of a routine voluntary approach.

By Simpson WM, Johnstone FD, Boyd FM, Goldberg DJ, Hart GJ, Gormley SM, *et al.*

No. 5

Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review.

By Koumounne OC, Gulliford MC, Chinn S, Sterne JAC, Burney PGJ.

No. 6

Assessing the costs of healthcare technologies in clinical trials.

A review by Johnston K, Buxton MJ, Jones DR, Fitzpatrick R.

No. 7

Cooperatives and their primary care emergency centres: organisation and impact.

By Hallam L, Henthorne K.

No. 8

Screening for cystic fibrosis.

A review by Murray J, Cuckle H, Taylor G, Littlewood J, Hewison J.

No. 9

A review of the use of health status measures in economic evaluation.

By Brazier J, Deverill M, Green C, Harper R, Booth A.

No. 10

Methods for the analysis of quality-of-life and survival data in health technology assessment.

A review by Billingham LJ, Abrams KR, Jones DR.

No. 11

Antenatal and neonatal haemoglobinopathy screening in the UK: review and economic analysis.

By Zeuner D, Ades AE, Karnon J, Brown J, Dezateux C, Anionwu EN.

No. 12

Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses.

A review by Moher D, Cook DJ, Jadad AR, Tugwell P, Moher M, Jones A, *et al.*

No. 13

'Early warning systems' for identifying new healthcare technologies.

By Robert G, Stevens A, Gabbay J.

No. 14

A systematic review of the role of human papillomavirus testing within a cervical screening programme.

By Cuzick J, Sasieni P, Davies P, Adams J, Normand C, Frater A, *et al.*

No. 15

Near patient testing in diabetes clinics: appraising the costs and outcomes.

By Grieve R, Beech R, Vincent J, Mazurkiewicz J.

No. 16

Positron emission tomography: establishing priorities for health technology assessment.

A review by Robert G, Milne R.

No. 17 (Pt 1)

The debridement of chronic wounds: a systematic review.

By Bradley M, Cullum N, Sheldon T.

No. 17 (Pt 2)

Systematic reviews of wound care management: (2) Dressings and topical agents used in the healing of chronic wounds.

By Bradley M, Cullum N, Nelson EA, Petticrew M, Sheldon T, Torgerson D.

No. 18

A systematic literature review of spiral and electron beam computed tomography: with particular reference to clinical applications in hepatic lesions, pulmonary embolus and coronary artery disease.

By Berry E, Kelly S, Hutton J, Harris KM, Roderick P, Boyce JC, *et al.*

No. 19

What role for statins? A review and economic model.

By Ebrahim S, Davey Smith G, McCabe C, Payne N, Pickin M, Sheldon TA, *et al.*

No. 20

Factors that limit the quality, number and progress of randomised controlled trials.

A review by Prescott RJ, Counsell CE, Gillespie WJ, Grant AM, Russell IT, Kiauka S, *et al.*

No. 21

Antimicrobial prophylaxis in total hip replacement: a systematic review.

By Glenny AM, Song F.

No. 22

Health promoting schools and health promotion in schools: two systematic reviews.

By Lister-Sharp D, Chapman S, Stewart-Brown S, Sowden A.

No. 23

Economic evaluation of a primary care-based education programme for patients with osteoarthritis of the knee.

A review by Lord J, Victor C, Littlejohns P, Ross FM, Axford JS.

Volume 4, 2000**No. 1**

The estimation of marginal time preference in a UK-wide sample (TEMPUS) project.

A review by Cairns JA, van der Pol MM.

No. 2

Geriatric rehabilitation following fractures in older people: a systematic review.

By Cameron I, Crotty M, Currie C, Finnegan T, Gillespie L, Gillespie W, *et al.*

No. 3

Screening for sickle cell disease and thalassaemia: a systematic review with supplementary research.

By Davies SC, Cronin E, Gill M, Greengross P, Hickman M, Normand C.

No. 4

Community provision of hearing aids and related audiology services.

A review by Reeves DJ, Alborz A, Hickson FS, Bamford JM.

No. 5

False-negative results in screening programmes: systematic review of impact and implications.

By Petticrew MP, Sowden AJ, Lister-Sharp D, Wright K.

No. 6

Costs and benefits of community postnatal support workers: a randomised controlled trial.

By Morrell CJ, Spiby H, Stewart P, Walters S, Morgan A.

No. 7

Implantable contraceptives (subdermal implants and hormonally impregnated intrauterine systems) versus other forms of reversible contraceptives: two systematic reviews to assess relative effectiveness, acceptability, tolerability and cost-effectiveness.

By French RS, Cowan FM, Mansour DJA, Morris S, Procter T, Hughes D, *et al.*

No. 8

An introduction to statistical methods for health technology assessment.

A review by White SJ, Ashby D, Brown PJ.

No. 9

Disease-modifying drugs for multiple sclerosis: a rapid and systematic review.

By Clegg A, Bryant J, Milne R.

No. 10

Publication and related biases.

A review by Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ.

No. 11

Cost and outcome implications of the organisation of vascular services.

By Michaels J, Brazier J, Palfreyman S, Shackley P, Slack R.

No. 12

Monitoring blood glucose control in diabetes mellitus: a systematic review.

By Coster S, Gulliford MC, Seed PT, Powrie JK, Swaminathan R.

No. 13

The effectiveness of domiciliary health visiting: a systematic review of international studies and a selective review of the British literature.

By Elkan R, Kendrick D, Hewitt M, Robinson JJA, Tolley K, Blair M, *et al.*

No. 14

The determinants of screening uptake and interventions for increasing uptake: a systematic review.

By Jepson R, Clegg A, Forbes C, Lewis R, Sowden A, Kleijnen J.

No. 15

The effectiveness and cost-effectiveness of prophylactic removal of wisdom teeth.

A rapid review by Song F, O'Meara S, Wilson P, Golder S, Kleijnen J.

No. 16

Ultrasound screening in pregnancy: a systematic review of the clinical effectiveness, cost-effectiveness and women's views.

By Bricker L, Garcia J, Henderson J, Mugford M, Neilson J, Roberts T, *et al.*

No. 17

A rapid and systematic review of the effectiveness and cost-effectiveness of the taxanes used in the treatment of advanced breast and ovarian cancer.

By Lister-Sharp D, McDonagh MS, Khan KS, Kleijnen J.

No. 18

Liquid-based cytology in cervical screening: a rapid and systematic review.

By Payne N, Chilcott J, McGoogan E.

No. 19

Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy and usual general practitioner care in the management of depression as well as mixed anxiety and depression in primary care.

By King M, Sibbald B, Ward E, Bower P, Lloyd M, Gabbay M, *et al.*

No. 20

Routine referral for radiography of patients presenting with low back pain: is patients' outcome influenced by GPs' referral for plain radiography?

By Kerry S, Hilton S, Patel S, Dundas D, Rink E, Lord J.

No. 21

Systematic reviews of wound care management: (3) antimicrobial agents for chronic wounds; (4) diabetic foot ulceration.

By O'Meara S, Cullum N, Majid M, Sheldon T.

No. 22

Using routine data to complement and enhance the results of randomised controlled trials.

By Lewsey JD, Leyland AH, Murray GD, Boddy FA.

No. 23

Coronary artery stents in the treatment of ischaemic heart disease: a rapid and systematic review.

By Meads C, Cummins C, Jolly K, Stevens A, Burls A, Hyde C.

No. 24

Outcome measures for adult critical care: a systematic review.

By Hayes JA, Black NA, Jenkinson C, Young JD, Rowan KM, Daly K, *et al.*

No. 25

A systematic review to evaluate the effectiveness of interventions to promote the initiation of breastfeeding.

By Fairbank L, O'Meara S, Renfrew MJ, Woolridge M, Sowden AJ, Lister-Sharp D.

No. 26

Implantable cardioverter defibrillators: arrhythmias. A rapid and systematic review.

By Parkes J, Bryant J, Milne R.

No. 27

Treatments for fatigue in multiple sclerosis: a rapid and systematic review.

By Brañas P, Jordan R, Fry-Smith A, Burls A, Hyde C.

No. 28

Early asthma prophylaxis, natural history, skeletal development and economy (EASE): a pilot randomised controlled trial.

By Baxter-Jones ADG, Helms PJ, Russell G, Grant A, Ross S, Cairns JA, *et al.*

No. 29

Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis.

By Marks D, Wonderling D, Thorogood M, Lambert H, Humphries SE, Neil HAW.

No. 30

A rapid and systematic review of the clinical effectiveness and cost-effectiveness of glycoprotein IIb/IIIa antagonists in the medical management of unstable angina.

By McDonagh MS, Bachmann LM, Golder S, Kleijnen J, ter Riet G.

No. 31

A randomised controlled trial of prehospital intravenous fluid replacement therapy in serious trauma.

By Turner J, Nicholl J, Webber L, Cox H, Dixon S, Yates D.

No. 32

Intrathecal pumps for giving opioids in chronic pain: a systematic review.

By Williams JE, Louw G, Towler G.

No. 33

Combination therapy (interferon alfa and ribavirin) in the treatment of chronic hepatitis C: a rapid and systematic review.

By Shepherd J, Waugh N, Hewitson P.

No. 34

A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies.

By MacLehose RR, Reeves BC, Harvey IM, Sheldon TA, Russell IT, Black AMS.

No. 35

Intravascular ultrasound-guided interventions in coronary artery disease: a systematic literature review, with decision-analytic modelling, of outcomes and cost-effectiveness.

By Berry E, Kelly S, Hutton J, Lindsay HSJ, Blaxill JM, Evans JA, *et al.*

No. 36

A randomised controlled trial to evaluate the effectiveness and cost-effectiveness of counselling patients with chronic depression.

By Simpson S, Corney R, Fitzgerald P, Beecham J.

No. 37

Systematic review of treatments for atopic eczema.

By Hoare C, Li Wan Po A, Williams H.

No. 38

Bayesian methods in health technology assessment: a review.

By Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR.

No. 39

The management of dyspepsia: a systematic review.

By Delaney B, Moayyedi P, Deeks J, Innes M, Soo S, Barton P, *et al.*

No. 40

A systematic review of treatments for severe psoriasis.

By Griffiths CEM, Clark CM, Chalmers RJG, Li Wan Po A, Williams HC.

Volume 5, 2001

No. 1

Clinical and cost-effectiveness of donepezil, rivastigmine and galantamine for Alzheimer's disease: a rapid and systematic review.

By Clegg A, Bryant J, Nicholson T, McIntyre L, De Broe S, Gerard K, *et al.*

No. 2

The clinical effectiveness and cost-effectiveness of riluzole for motor neurone disease: a rapid and systematic review.

By Stewart A, Sandercock J, Bryan S, Hyde C, Barton PM, Fry-Smith A, *et al.*

No. 3

Equity and the economic evaluation of healthcare.

By Sassi F, Archard L, Le Grand J.

No. 4

Quality-of-life measures in chronic diseases of childhood.

By Eiser C, Morse R.

No. 5

Eliciting public preferences for healthcare: a systematic review of techniques.

By Ryan M, Scott DA, Reeves C, Bate A, van Teijlingen ER, Russell EM, *et al.*

No. 6

General health status measures for people with cognitive impairment: learning disability and acquired brain injury.

By Riemsma RP, Forbes CA, Glanville JM, Eastwood AJ, Kleijnen J.

No. 7

An assessment of screening strategies for fragile X syndrome in the UK.

By Pembrey ME, Barnicoat AJ, Carmichael B, Bobrow M, Turner G.

No. 8

Issues in methodological research: perspectives from researchers and commissioners.

By Lilford RJ, Richardson A, Stevens A, Fitzpatrick R, Edwards S, Rock F, *et al.*

No. 9

Systematic reviews of wound care management: (5) beds; (6) compression; (7) laser therapy, therapeutic ultrasound, electrotherapy and electromagnetic therapy.

By Cullum N, Nelson EA, Flemming K, Sheldon T.

No. 10

Effects of educational and psychosocial interventions for adolescents with diabetes mellitus: a systematic review.

By Hampson SE, Skinner TC, Hart J, Storey L, Gage H, Foxcroft D, *et al.*

No. 11

Effectiveness of autologous chondrocyte transplantation for hyaline cartilage defects in knees: a rapid and systematic review.

By Jobanputra P, Parry D, Fry-Smith A, Burls A.

No. 12

Statistical assessment of the learning curves of health technologies.

By Ramsay CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF, Russell IT.

No. 13

The effectiveness and cost-effectiveness of temozolomide for the treatment of recurrent malignant glioma: a rapid and systematic review.

By Dinnes J, Cave C, Huang S, Major K, Milne R.

No. 14

A rapid and systematic review of the clinical effectiveness and cost-effectiveness of debriding agents in treating surgical wounds healing by secondary intention.

By Lewis R, Whiting P, ter Riet G, O'Meara S, Glanville J.

No. 15

Home treatment for mental health problems: a systematic review.

By Burns T, Knapp M, Catty J, Healey A, Henderson J, Watt H, *et al.*

No. 16

How to develop cost-conscious guidelines.

By Eccles M, Mason J.

No. 17

The role of specialist nurses in multiple sclerosis: a rapid and systematic review.

By De Broe S, Christopher F, Waugh N.

No. 18

A rapid and systematic review of the clinical effectiveness and cost-effectiveness of orlistat in the management of obesity.

By O'Meara S, Riemsma R, Shirran L, Mather L, ter Riet G.

No. 19

The clinical effectiveness and cost-effectiveness of pioglitazone for type 2 diabetes mellitus: a rapid and systematic review.

By Chilcott J, Wight J, Lloyd Jones M, Tappenden P.

No. 20

Extended scope of nursing practice: a multicentre randomised controlled trial of appropriately trained nurses and preregistration house officers in preoperative assessment in elective general surgery.

By Kinley H, Czoski-Murray C, George S, McCabe C, Primrose J, Reilly C, *et al.*

No. 21

Systematic reviews of the effectiveness of day care for people with severe mental disorders: (1) Acute day hospital versus admission; (2) Vocational rehabilitation; (3) Day hospital versus outpatient care.

By Marshall M, Crowther R, Almaraz-Serrano A, Creed F, Sledge W, Kluiters H, *et al.*

No. 22

The measurement and monitoring of surgical adverse events.

By Bruce J, Russell EM, Mollison J, Krukowski ZH.

No. 23

Action research: a systematic review and guidance for assessment.

By Waterman H, Tillen D, Dickson R, de Koning K.

No. 24

A rapid and systematic review of the clinical effectiveness and cost-effectiveness of gemcitabine for the treatment of pancreatic cancer.

By Ward S, Morris E, Bansback N, Calvert N, Crellin A, Forman D, *et al.*

No. 25

A rapid and systematic review of the evidence for the clinical effectiveness and cost-effectiveness of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer.

By Lloyd Jones M, Hummel S, Bansback N, Orr B, Seymour M.

No. 26

Comparison of the effectiveness of inhaler devices in asthma and chronic obstructive airways disease: a systematic review of the literature.

By Brocklebank D, Ram F, Wright J, Barry P, Cates C, Davies L, *et al.*

No. 27

The cost-effectiveness of magnetic resonance imaging for investigation of the knee joint.

By Bryan S, Weatherburn G, Bungay H, Hatrick C, Salas C, Parry D, *et al.*

No. 28

A rapid and systematic review of the clinical effectiveness and cost-effectiveness of topotecan for ovarian cancer.

By Forbes C, Shirran L, Bagnall A-M, Duffy S, ter Riet G.

No. 29

Superseded by a report published in a later volume.

No. 30

The role of radiography in primary care patients with low back pain of at least 6 weeks duration: a randomised (unblinded) controlled trial.

By Kendrick D, Fielding K, Bentley E, Miller P, Kerslake R, Pringle M.

No. 31

Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients.

By McColl E, Jacoby A, Thomas L, Soutter J, Bamford C, Steen N, *et al.*

No. 32

A rapid and systematic review of the clinical effectiveness and cost-effectiveness of paclitaxel, docetaxel, gemcitabine and vinorelbine in non-small-cell lung cancer.

By Clegg A, Scott DA, Sidhu M, Hewitson P, Waugh N.

No. 33

Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives.

By Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G.

No. 34

Depot antipsychotic medication in the treatment of patients with schizophrenia: (1) Meta-review; (2) Patient and nurse attitudes.

By David AS, Adams C.

No. 35

A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression.

By Churchill R, Hunot V, Corney R, Knapp M, McGuire H, Tylee A, *et al.*

No. 36

Cost analysis of child health surveillance.

By Sanderson D, Wright D, Acton C, Duree D.

Volume 6, 2002**No. 1**

A study of the methods used to select review criteria for clinical audit.

By Hearnshaw H, Harker R, Cheater F, Baker R, Grimshaw G.

No. 2

Fludarabine as second-line therapy for B cell chronic lymphocytic leukaemia: a technology assessment.

By Hyde C, Wake B, Bryan S, Barton P, Fry-Smith A, Davenport C, *et al.*

No. 3

Rituximab as third-line treatment for refractory or recurrent Stage III or IV follicular non-Hodgkin's lymphoma: a systematic review and economic evaluation.

By Wake B, Hyde C, Bryan S, Barton P, Song F, Fry-Smith A, *et al.*

No. 4

A systematic review of discharge arrangements for older people.

By Parker SG, Peet SM, McPherson A, Cannaby AM, Baker R, Wilson A, *et al.*

No. 5

The clinical effectiveness and cost-effectiveness of inhaler devices used in the routine management of chronic asthma in older children: a systematic review and economic evaluation.

By Peters J, Stevenson M, Beverley C, Lim J, Smith S.

No. 6

The clinical effectiveness and cost-effectiveness of sibutramine in the management of obesity: a technology assessment.

By O'Meara S, Riemsma R, Shirran L, Mather L, ter Riet G.

No. 7

The cost-effectiveness of magnetic resonance angiography for carotid artery stenosis and peripheral vascular disease: a systematic review.

By Berry E, Kelly S, Westwood ME, Davies LM, Gough MJ, Bamford JM, *et al.*

No. 8

Promoting physical activity in South Asian Muslim women through 'exercise on prescription'.

By Carroll B, Ali N, Azam N.

No. 9

Zanamivir for the treatment of influenza in adults: a systematic review and economic evaluation.

By Burls A, Clark W, Stewart T, Preston C, Bryan S, Jefferson T, *et al.*

No. 10

A review of the natural history and epidemiology of multiple sclerosis: implications for resource allocation and health economic models.

By Richards RG, Sampson FC, Beard SM, Tappenden P.

No. 11

Screening for gestational diabetes: a systematic review and economic evaluation.

By Scott DA, Loveman E, McIntyre L, Waugh N.

No. 12

The clinical effectiveness and cost-effectiveness of surgery for people with morbid obesity: a systematic review and economic evaluation.

By Clegg AJ, Colquitt J, Sidhu MK, Royle P, Loveman E, Walker A.

No. 13

The clinical effectiveness of trastuzumab for breast cancer: a systematic review.

By Lewis R, Bagnall A-M, Forbes C, Shirran E, Duffy S, Kleijnen J, *et al.*

No. 14

The clinical effectiveness and cost-effectiveness of vinorelbine for breast cancer: a systematic review and economic evaluation.

By Lewis R, Bagnall A-M, King S, Woolcott N, Forbes C, Shirran L, *et al.*

No. 15

A systematic review of the effectiveness and cost-effectiveness of metal-on-metal hip resurfacing arthroplasty for treatment of hip disease.

By Vale L, Wyness L, McCormack K, McKenzie L, Brazzelli M, Stearns SC.

No. 16

The clinical effectiveness and cost-effectiveness of bupropion and nicotine replacement therapy for smoking cessation: a systematic review and economic evaluation.

By Woolcott NF, Jones L, Forbes CA, Mather LC, Sowden AJ, Song FJ, *et al.*

No. 17

A systematic review of effectiveness and economic evaluation of new drug treatments for juvenile idiopathic arthritis: etanercept.

By Cummins C, Connock M, Fry-Smith A, Burls A.

No. 18

Clinical effectiveness and cost-effectiveness of growth hormone in children: a systematic review and economic evaluation.

By Bryant J, Cave C, Mihaylova B, Chase D, McIntyre L, Gerard K, *et al.*

No. 19

Clinical effectiveness and cost-effectiveness of growth hormone in adults in relation to impact on quality of life: a systematic review and economic evaluation.

By Bryant J, Loveman E, Chase D, Mihaylova B, Cave C, Gerard K, *et al.*

No. 20

Clinical medication review by a pharmacist of patients on repeat prescriptions in general practice: a randomised controlled trial.

By Zermansky AG, Petty DR, Raynor DK, Lowe CJ, Freemantle N, Vail A.

No. 21

The effectiveness of infliximab and etanercept for the treatment of rheumatoid arthritis: a systematic review and economic evaluation.

By Jobanputra P, Barton P, Bryan S, Burls A.

No. 22

A systematic review and economic evaluation of computerised cognitive behaviour therapy for depression and anxiety.

By Kaltenthaler E, Shackley P, Stevens K, Beverley C, Parry G, Chilcott J.

No. 23

A systematic review and economic evaluation of pegylated liposomal doxorubicin hydrochloride for ovarian cancer.

By Forbes C, Wilby J, Richardson G, Sculpher M, Mather L, Riemsma R.

No. 24

A systematic review of the effectiveness of interventions based on a stages-of-change approach to promote individual behaviour change.

By Riemsma RP, Pattenden J, Bridle C, Sowden AJ, Mather L, Watt IS, *et al.*

No. 25

A systematic review update of the clinical effectiveness and cost-effectiveness of glycoprotein IIb/IIIa antagonists.

By Robinson M, Ginnelly L, Sculpher M, Jones L, Riemsma R, Palmer S, *et al.*

No. 26

A systematic review of the effectiveness, cost-effectiveness and barriers to implementation of thrombolytic and neuroprotective therapy for acute ischaemic stroke in the NHS.

By Sandercock P, Berge E, Dennis M, Forbes J, Hand P, Kwan J, *et al.*

No. 27

A randomised controlled crossover trial of nurse practitioner versus doctor-led outpatient care in a bronchiectasis clinic.

By Caine N, Sharples LD, Hollingworth W, French J, Keogan M, Exley A, *et al.*

No. 28

Clinical effectiveness and cost – consequences of selective serotonin reuptake inhibitors in the treatment of sex offenders.

By Adi Y, Ashcroft D, Browne K, Beech A, Fry-Smith A, Hyde C.

No. 29

Treatment of established osteoporosis: a systematic review and cost-utility analysis.

By Kanis JA, Brazier JE, Stevenson M, Calvert NW, Lloyd Jones M.

No. 30

Which anaesthetic agents are cost-effective in day surgery? Literature review, national survey of practice and randomised controlled trial.

By Elliott RA Payne K, Moore JK, Davies LM, Harper NJN, St Leger AS, *et al.*

No. 31

Screening for hepatitis C among injecting drug users and in genitourinary medicine clinics: systematic reviews of effectiveness, modelling study and national survey of current practice.

By Stein K, Dalziel K, Walker A, McIntyre L, Jenkins B, Horne J, *et al.*

No. 32

The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature.

By Crow R, Gage H, Hampson S, Hart J, Kimber A, Storey L, *et al.*

No. 33

The effectiveness and cost-effectiveness of imatinib in chronic myeloid leukaemia: a systematic review.

By Garside R, Round A, Dalziel K, Stein K, Royle R.

No. 34

A comparative study of hypertonic saline, daily and alternate-day rhDNase in children with cystic fibrosis.

By Suri R, Wallis C, Bush A, Thompson S, Normand C, Flather M, *et al.*

No. 35

A systematic review of the costs and effectiveness of different models of paediatric home care.

By Parker G, Bhakta P, Lovett CA, Paisley S, Olsen R, Turner D, *et al.*

Volume 7, 2003

No. 1

How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study.

By Egger M, Jüni P, Bartlett C, Holenstein F, Sterne J.

No. 2

Systematic review of the effectiveness and cost-effectiveness, and economic evaluation, of home versus hospital or satellite unit haemodialysis for people with end-stage renal failure.

By Mowatt G, Vale L, Perez J, Wyness L, Fraser C, MacLeod A, *et al.*

No. 3

Systematic review and economic evaluation of the effectiveness of infliximab for the treatment of Crohn's disease.

By Clark W, Raftery J, Barton P, Song F, Fry-Smith A, Burls A.

No. 4

A review of the clinical effectiveness and cost-effectiveness of routine anti-D prophylaxis for pregnant women who are rhesus negative.

By Chilcott J, Lloyd Jones M, Wight J, Forman K, Wray J, Beverley C, *et al.*

No. 5

Systematic review and evaluation of the use of tumour markers in paediatric oncology: Ewing's sarcoma and neuroblastoma.

By Riley RD, Burchill SA, Abrams KR, Heney D, Lambert PC, Jones DR, *et al.*

No. 6

The cost-effectiveness of screening for *Helicobacter pylori* to reduce mortality and morbidity from gastric cancer and peptic ulcer disease: a discrete-event simulation model.

By Roderick P, Davies R, Raftery J, Crabbe D, Pearce R, Bhandari P, *et al.*

No. 7

The clinical effectiveness and cost-effectiveness of routine dental checks: a systematic review and economic evaluation.

By Davenport C, Elley K, Salas C, Taylor-Weetman CL, Fry-Smith A, Bryan S, *et al.*

No. 8

A multicentre randomised controlled trial assessing the costs and benefits of using structured information and analysis of women's preferences in the management of menorrhagia.

By Kennedy ADM, Sculpher MJ, Coulter A, Dwyer N, Rees M, Horsley S, *et al.*

No. 9

Clinical effectiveness and cost-utility of photodynamic therapy for wet age-related macular degeneration: a systematic review and economic evaluation.

By Meads C, Salas C, Roberts T, Moore D, Fry-Smith A, Hyde C.

No. 10

Evaluation of molecular tests for prenatal diagnosis of chromosome abnormalities.

By Grimshaw GM, Szczepura A, Hultén M, MacDonald F, Nevin NC, Sutton F, *et al.*

No. 11

First and second trimester antenatal screening for Down's syndrome: the results of the Serum, Urine and Ultrasound Screening Study (SURUSS).

By Wald NJ, Rodeck C, Hackshaw AK, Walters J, Chitty L, Mackinson AM.

No. 12

The effectiveness and cost-effectiveness of ultrasound locating devices for central venous access: a systematic review and economic evaluation.

By Calvert N, Hind D, McWilliams RG, Thomas SM, Beverley C, Davidson A.

No. 13

A systematic review of atypical antipsychotics in schizophrenia.

By Bagnall A-M, Jones L, Lewis R, Ginnelly L, Glanville J, Torgerson D, *et al.*

No. 14

Prostate Testing for Cancer and Treatment (ProtecT) feasibility study.

By Donovan J, Hamdy F, Neal D, Peters T, Oliver S, Brindle L, *et al.*

No. 15

Early thrombolysis for the treatment of acute myocardial infarction: a systematic review and economic evaluation.

By Boland A, Dundar Y, Bagust A, Haycox A, Hill R, Mujica Mota R, *et al.*

No. 16

Screening for fragile X syndrome: a literature review and modelling.

By Song FJ, Barton P, Sleightholme V, Yao GL, Fry-Smith A.

No. 17

Systematic review of endoscopic sinus surgery for nasal polyps.

By Dalziel K, Stein K, Round A, Garside R, Royle P.

No. 18

Towards efficient guidelines: how to monitor guideline use in primary care.

By Hutchinson A, McIntosh A, Cox S, Gilbert C.

No. 19

Effectiveness and cost-effectiveness of acute hospital-based spinal cord injuries services: systematic review.

By Bagnall A-M, Jones L, Richardson G, Duffy S, Riemsma R.

No. 20

Prioritisation of health technology assessment. The PATHS model: methods and case studies.

By Townsend J, Buxton M, Harper G.

No. 21

Systematic review of the clinical effectiveness and cost-effectiveness of tension-free vaginal tape for treatment of urinary stress incontinence.

By Cody J, Wyness L, Wallace S, Glazener C, Kilonzo M, Stearns S, *et al.*

No. 22

The clinical and cost-effectiveness of patient education models for diabetes: a systematic review and economic evaluation.

By Loveman E, Cave C, Green C, Royle P, Dunn N, Waugh N.

No. 23

The role of modelling in prioritising and planning clinical trials.

By Chilcott J, Brennan A, Booth A, Karnon J, Tappenden P.

No. 24

Cost-benefit evaluation of routine influenza immunisation in people 65-74 years of age.

By Allsup S, Gosney M, Haycox A, Regan M.

No. 25

The clinical and cost-effectiveness of pulsatile machine perfusion versus cold storage of kidneys for transplantation retrieved from heart-beating and non-heart-beating donors.

By Wight J, Chilcott J, Holmes M, Brewer N.

No. 26

Can randomised trials rely on existing electronic data? A feasibility study to explore the value of routine data in health technology assessment.

By Williams JG, Cheung WY, Cohen DR, Hutchings HA, Longo MF, Russell IT.

No. 27

Evaluating non-randomised intervention studies.

By Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovich C, Song F, *et al.*

No. 28

A randomised controlled trial to assess the impact of a package comprising a patient-orientated, evidence-based self-help guidebook and patient-centred consultations on disease management and satisfaction in inflammatory bowel disease.

By Kennedy A, Nelson E, Reeves D, Richardson G, Roberts C, Robinson A, *et al.*

No. 29

The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review.

By Dinnes J, Loveman E, McIntyre L, Waugh N.

No. 30

The value of digital imaging in diabetic retinopathy.

By Sharp PF, Olson J, Strachan F, Hipwell J, Ludbrook A, O'Donnell M, *et al.*

No. 31

Lowering blood pressure to prevent myocardial infarction and stroke: a new preventive strategy.

By Law M, Wald N, Morris J.

No. 32

Clinical and cost-effectiveness of capecitabine and tegafur with uracil for the treatment of metastatic colorectal cancer: systematic review and economic evaluation.

By Ward S, Kaltenthaler E, Cowan J, Brewer N.

No. 33

Clinical and cost-effectiveness of new and emerging technologies for early localised prostate cancer: a systematic review.

By Hummel S, Paisley S, Morgan A, Currie E, Brewer N.

No. 34

Literature searching for clinical and cost-effectiveness studies used in health technology assessment reports carried out for the National Institute for Clinical Excellence appraisal system.

By Royle P, Waugh N.

No. 35

Systematic review and economic decision modelling for the prevention and treatment of influenza A and B.

By Turner D, Wailoo A, Nicholson K, Cooper N, Sutton A, Abrams K.

No. 36

A randomised controlled trial to evaluate the clinical and cost-effectiveness of Hickman line insertions in adult cancer patients by nurses.

By Boland A, Haycox A, Bagust A, Fitzsimmons L.

No. 37

Redesigning postnatal care: a randomised controlled trial of protocol-based midwifery-led care focused on individual women's physical and psychological health needs.

By MacArthur C, Winter HR, Bick DE, Lilford RJ, Lancashire RJ, Knowles H, *et al.*

No. 38

Estimating implied rates of discount in healthcare decision-making.

By West RR, McNabb R, Thompson AGH, Sheldon TA, Grimley Evans J.

No. 39

Systematic review of isolation policies in the hospital management of methicillin-resistant *Staphylococcus aureus*: a review of the literature with epidemiological and economic modelling.

By Cooper BS, Stone SP, Kibbler CC, Cookson BD, Roberts JA, Medley GF, *et al.*

No. 40

Treatments for spasticity and pain in multiple sclerosis: a systematic review.

By Beard S, Hunn A, Wight J.

No. 41

The inclusion of reports of randomised trials published in languages other than English in systematic reviews.

By Moher D, Pham B, Lawson ML, Klassen TP.

No. 42

The impact of screening on future health-promoting behaviours and health beliefs: a systematic review.

By Bankhead CR, Brett J, Bukach C, Webster P, Stewart-Brown S, Munafo M, *et al.*

Volume 8, 2004

No. 1

What is the best imaging strategy for acute stroke?

By Wardlaw JM, Keir SL, Seymour J, Lewis S, Sandercock PAG, Dennis MS, *et al.*

No. 2

Systematic review and modelling of the investigation of acute and chronic chest pain presenting in primary care.

By Mant J, McManus RJ, Oakes RAL, Delaney BC, Barton PM, Deeks JJ, *et al.*

No. 3

The effectiveness and cost-effectiveness of microwave and thermal balloon endometrial ablation for heavy menstrual bleeding: a systematic review and economic modelling.

By Garside R, Stein K, Wyatt K, Round A, Price A.

No. 4

A systematic review of the role of bisphosphonates in metastatic disease.

By Ross JR, Saunders Y, Edmonds PM, Patel S, Wonderling D, Normand C, *et al.*

No. 5

Systematic review of the clinical effectiveness and cost-effectiveness of capecitabine (Xeloda®) for locally advanced and/or metastatic breast cancer.

By Jones L, Hawkins N, Westwood M, Wright K, Richardson G, Riemsma R.

No. 6

Effectiveness and efficiency of guideline dissemination and implementation strategies.

By Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, *et al.*

No. 7

Clinical effectiveness and costs of the Sugarbaker procedure for the treatment of pseudomyxoma peritonei.

By Bryant J, Clegg AJ, Sidhu MK, Brodin H, Royle P, Davidson P.

No. 8

Psychological treatment for insomnia in the regulation of long-term hypnotic drug use.

By Morgan K, Dixon S, Mathers N, Thompson J, Tomeny M.

No. 9

Improving the evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based measure of outcome.

By Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ.

No. 10

A systematic review and economic evaluation of magnetic resonance cholangiopancreatography compared with diagnostic endoscopic retrograde cholangiopancreatography.

By Kaltenthaler E, Bravo Vergel Y, Chilcott J, Thomas S, Blakeborough T, Walters SJ, *et al.*

No. 11

The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis.

By Barton P, Jobanputra P, Wilson J, Bryan S, Burls A.

No. 12

Clinical effectiveness and cost-effectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: a systematic review.

By Pandor A, Eastham J, Beverley C, Chilcott J, Paisley S.

No. 13

Clinical effectiveness and cost-effectiveness of pioglitazone and rosiglitazone in the treatment of type 2 diabetes: a systematic review and economic evaluation.

By Czoski-Murray C, Warren E, Chilcott J, Beverley C, Psyllaki MA, Cowan J.

No. 14

Routine examination of the newborn: the EMREN study. Evaluation of an extension of the midwife role including a randomised controlled trial of appropriately trained midwives and paediatric senior house officers.

By Townsend J, Wolke D, Hayes J, Davé S, Rogers C, Bloomfield L, *et al.*

No. 15

Involving consumers in research and development agenda setting for the NHS: developing an evidence-based approach.

By Oliver S, Clarke-Jones L, Rees R, Milne R, Buchanan P, Gabbay J, *et al.*

No. 16

A multi-centre randomised controlled trial of minimally invasive direct coronary bypass grafting versus percutaneous transluminal coronary angioplasty with stenting for proximal stenosis of the left anterior descending coronary artery.

By Reeves BC, Angelini GD, Bryan AJ, Taylor FC, Cripps T, Spyt TJ, *et al.*

No. 17

Does early magnetic resonance imaging influence management or improve outcome in patients referred to secondary care with low back pain? A pragmatic randomised controlled trial.

By Gilbert FJ, Grant AM, Gillan MGC, Vale L, Scott NW, Campbell MK, *et al.*

No. 18

The clinical and cost-effectiveness of anakinra for the treatment of rheumatoid arthritis in adults: a systematic review and economic analysis.

By Clark W, Jobanputra P, Barton P, Burls A.

No. 19

A rapid and systematic review and economic evaluation of the clinical and cost-effectiveness of newer drugs for treatment of mania associated with bipolar affective disorder.

By Bridle C, Palmer S, Bagnall A-M, Darba J, Duffy S, Sculpher M, *et al.*

No. 20

Liquid-based cytology in cervical screening: an updated rapid and systematic review and economic analysis.

By Karnon J, Peters J, Platt J, Chilcott J, McGoogan E, Brewer N.

No. 21

Systematic review of the long-term effects and economic consequences of treatments for obesity and implications for health improvement.

By Avenell A, Broom J, Brown TJ, Poobalan A, Aucott L, Stearns SC, *et al.*

No. 22

Autoantibody testing in children with newly diagnosed type 1 diabetes mellitus.

By Dretzke J, Cummins C, Sandercock J, Fry-Smith A, Barrett T, Burls A.

No. 23

Clinical effectiveness and cost-effectiveness of prehospital intravenous fluids in trauma patients.

By Dretzke J, Sandercock J, Bayliss S, Burls A.

No. 24

Newer hypnotic drugs for the short-term management of insomnia: a systematic review and economic evaluation.

By Dündar Y, Boland A, Strobl J, Dodd S, Haycox A, Bagust A, *et al.*

No. 25

Development and validation of methods for assessing the quality of diagnostic accuracy studies.

By Whiting P, Rutjes AWS, Dinnes J, Reitsma JB, Bossuyt PMM, Kleijnen J.

No. 26

EVALUATE hysterectomy trial: a multicentre randomised trial comparing abdominal, vaginal and laparoscopic methods of hysterectomy.

By Garry R, Fountain J, Brown J, Manca A, Mason S, Sculpher M, *et al.*

No. 27

Methods for expected value of information analysis in complex health economic models: developments on the health economics of interferon- β and glatiramer acetate for multiple sclerosis.

By Tappenden P, Chilcott JB, Eggington S, Oakley J, McCabe C.

No. 28

Effectiveness and cost-effectiveness of imatinib for first-line treatment of chronic myeloid leukaemia in chronic phase: a systematic review and economic analysis.

By Dalziel K, Round A, Stein K, Garside R, Price A.

No. 29

VenUS I: a randomised controlled trial of two types of bandage for treating venous leg ulcers.

By Iglesias C, Nelson EA, Cullum NA, Torgerson DJ, on behalf of the VenUS Team.

No. 30

Systematic review of the effectiveness and cost-effectiveness, and economic evaluation, of myocardial perfusion scintigraphy for the diagnosis and management of angina and myocardial infarction.

By Mowatt G, Vale L, Brazzelli M, Hernandez R, Murray A, Scott N, *et al.*

No. 31

A pilot study on the use of decision theory and value of information analysis as part of the NHS Health Technology Assessment programme.

By Claxton K, Ginnelly L, Sculpher M, Philips Z, Palmer S.

No. 32

The Social Support and Family Health Study: a randomised controlled trial and economic evaluation of two alternative forms of postnatal support for mothers living in disadvantaged inner-city areas.

By Wiggins M, Oakley A, Roberts I, Turner H, Rajan L, Austerberry H, *et al.*

No. 33

Psychosocial aspects of genetic screening of pregnant women and newborns: a systematic review.

By Green JM, Hewison J, Bekker HL, Bryant, Cuckle HS.

No. 34

Evaluation of abnormal uterine bleeding: comparison of three outpatient procedures within cohorts defined by age and menopausal status.

By Critchley HOD, Warner P, Lee AJ, Brechin S, Guise J, Graham B.

No. 35

Coronary artery stents: a rapid systematic review and economic evaluation.

By Hill R, Bagust A, Bakhai A, Dickson R, Dündar Y, Haycox A, *et al.*

No. 36

Review of guidelines for good practice in decision-analytic modelling in health technology assessment.

By Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, *et al.*

No. 37

Rituximab (MabThera®) for aggressive non-Hodgkin's lymphoma: systematic review and economic evaluation.

By Knight C, Hind D, Brewer N, Abbott V.

No. 38

Clinical effectiveness and cost-effectiveness of clopidogrel and modified-release dipyridamole in the secondary prevention of occlusive vascular events: a systematic review and economic evaluation.

By Jones L, Griffin S, Palmer S, Main C, Orton V, Sculpher M, *et al.*

No. 39

Pegylated interferon α -2a and -2b in combination with ribavirin in the treatment of chronic hepatitis C: a systematic review and economic evaluation.

By Shepherd J, Brodin H, Cave C, Waugh N, Price A, Gabbay J.

No. 40

Clopidogrel used in combination with aspirin compared with aspirin alone in the treatment of non-ST-segment-elevation acute coronary syndromes: a systematic review and economic evaluation.

By Main C, Palmer S, Griffin S, Jones L, Orton V, Sculpher M, *et al.*

No. 41

Provision, uptake and cost of cardiac rehabilitation programmes: improving services to under-represented groups.

By Beswick AD, Rees K, Griesch I, Taylor FC, Burke M, West RR, *et al.*

No. 42

Involving South Asian patients in clinical trials.

By Hussain-Gambles M, Leese B, Atkin K, Brown J, Mason S, Tovey P.

No. 43

Clinical and cost-effectiveness of continuous subcutaneous insulin infusion for diabetes.

By Colquitt JL, Green C, Sidhu MK, Hartwell D, Waugh N.

No. 44

Identification and assessment of ongoing trials in health technology assessment reviews.

By Song FJ, Fry-Smith A, Davenport C, Bayliss S, Adi Y, Wilson JS, *et al.*

No. 45

Systematic review and economic evaluation of a long-acting insulin analogue, insulin glargine

By Warren E, Weatherley-Jones E, Chilcott J, Beverley C.

No. 46

Supplementation of a home-based exercise programme with a class-based programme for people with osteoarthritis of the knees: a randomised controlled trial and health economic analysis.

By McCarthy CJ, Mills PM, Pullen R, Richardson G, Hawkins N, Roberts CR, *et al.*

No. 47

Clinical and cost-effectiveness of once-daily versus more frequent use of same potency topical corticosteroids for atopic eczema: a systematic review and economic evaluation.

By Green C, Colquitt JL, Kirby J, Davidson P, Payne E.

No. 48

Acupuncture of chronic headache disorders in primary care: randomised controlled trial and economic analysis.

By Vickers AJ, Rees RW, Zollman CE, McCarney R, Smith CM, Ellis N, *et al.*

No. 49

Generalisability in economic evaluation studies in healthcare: a review and case studies.

By Sculpher MJ, Pang FS, Manca A, Drummond MF, Golder S, Urdahl H, *et al.*

No. 50

Virtual outreach: a randomised controlled trial and economic evaluation of joint teleconferenced medical consultations.

By Wallace P, Barber J, Clayton W, Currell R, Fleming K, Garner P, *et al.*

Volume 9, 2005

No. 1

Randomised controlled multiple treatment comparison to provide a cost-effectiveness rationale for the selection of antimicrobial therapy in acne.

By Ozolins M, Eady EA, Avery A, Cunliffe WJ, O'Neill C, Simpson NB, *et al.*

No. 2

Do the findings of case series studies vary significantly according to methodological characteristics?

By Dalziel K, Round A, Stein K, Garside R, Castelnovo E, Payne L.

No. 3

Improving the referral process for familial breast cancer genetic counselling: findings of three randomised controlled trials of two interventions.

By Wilson BJ, Torrance N, Mollison J, Wordsworth S, Gray JR, Haites NE, *et al.*

No. 4

Randomised evaluation of alternative electrosurgical modalities to treat bladder outflow obstruction in men with benign prostatic hyperplasia.

By Fowler C, McAllister W, Plail R, Karim O, Yang Q.

No. 5

A pragmatic randomised controlled trial of the cost-effectiveness of palliative therapies for patients with inoperable oesophageal cancer.

By Shenfine J, McNamee P, Steen N, Bond J, Griffin SM.

No. 6

Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography.

By Taylor P, Champness J, Given-Wilson R, Johnston K, Potts H.

No. 7

Issues in data monitoring and interim analysis of trials.

By Grant AM, Altman DG, Babiker AB, Campbell MK, Clemens FJ, Darbyshire JH, *et al.*

No. 8

Lay public's understanding of equipoise and randomisation in randomised controlled trials.

By Robinson EJ, Kerr CEP, Stevens AJ, Lilford RJ, Brauholtz DA, Edwards SJ, *et al.*

No. 9

Clinical and cost-effectiveness of electroconvulsive therapy for depressive illness, schizophrenia, catatonia and mania: systematic reviews and economic modelling studies.

By Greenhalgh J, Knight C, Hind D, Beverley C, Walters S.

No. 10

Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology.

By Smith SC, Lamping DL, Banerjee S, Harwood R, Foley B, Smith P, *et al.*

No. 11

Clinical effectiveness and cost-effectiveness of drotrecogin alfa (activated) (Xigris®) for the treatment of severe sepsis in adults: a systematic review and economic evaluation.

By Green C, Dinnes J, Takeda A, Shepherd J, Hartwell D, Cave C, *et al.*

No. 12

A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy.

By Dinnes J, Deeks J, Kirby J, Roderick P.

No. 13

Cervical screening programmes: can automation help? Evidence from systematic reviews, an economic analysis and a simulation modelling exercise applied to the UK.

By Willis BH, Barton P, Pearmain P, Bryan S, Hyde C.

No. 14

Laparoscopic surgery for inguinal hernia repair: systematic review of effectiveness and economic evaluation.

By McCormack K, Wake B, Perez J, Fraser C, Cook J, McIntosh E, *et al.*

No. 15

Clinical effectiveness, tolerability and cost-effectiveness of newer drugs for epilepsy in adults: a systematic review and economic evaluation.

By Wilby J, Kainth A, Hawkins N, Epstein D, McIntosh H, McDaid C, *et al.*

No. 16

A randomised controlled trial to compare the cost-effectiveness of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine.

By Peveler R, Kendrick T, Buxton M, Longworth L, Baldwin D, Moore M, *et al.*

No. 17

Clinical effectiveness and cost-effectiveness of immediate angioplasty for acute myocardial infarction: systematic review and economic evaluation.

By Hartwell D, Colquitt J, Loveman E, Clegg AJ, Brodin H, Waugh N, *et al.*

No. 18

A randomised controlled comparison of alternative strategies in stroke care.

By Kalra L, Evans A, Perez I, Knapp M, Swift C, Donaldson N.

No. 19

The investigation and analysis of critical incidents and adverse events in healthcare.

By Woloshynowych M, Rogers S, Taylor-Adams S, Vincent C.

No. 20

Potential use of routine databases in health technology assessment.

By Raftery J, Roderick P, Stevens A.

No. 21

Clinical and cost-effectiveness of newer immunosuppressive regimens in renal transplantation: a systematic review and modelling study.

By Woodroffe R, Yao GL, Meads C, Bayliss S, Ready A, Raftery J, *et al.*

No. 22

A systematic review and economic evaluation of alendronate, etidronate, risedronate, raloxifene and teriparatide for the prevention and treatment of postmenopausal osteoporosis.

By Stevenson M, Lloyd Jones M, De Nigris E, Brewer N, Davis S, Oakley J.

No. 23

A systematic review to examine the impact of psycho-educational interventions on health outcomes and costs in adults and children with difficult asthma.

By Smith JR, Mugford M, Holland R, Candy B, Noble MJ, Harrison BDW, *et al.*

No. 24

An evaluation of the costs, effectiveness and quality of renal replacement therapy provision in renal satellite units in England and Wales.

By Roderick P, Nicholson T, Armitage A, Mehta R, Mullee M, Gerard K, *et al.*

No. 25

Imatinib for the treatment of patients with unresectable and/or metastatic gastrointestinal stromal tumours: systematic review and economic evaluation.

By Wilson J, Connock M, Song F, Yao G, Fry-Smith A, Raftery J, *et al.*

No. 26

Indirect comparisons of competing interventions.

By Glenny AM, Altman DG, Song F, Sakarovich C, Deeks JJ, D'Amico R, *et al.*

No. 27

Cost-effectiveness of alternative strategies for the initial medical management of non-ST elevation acute coronary syndrome: systematic review and decision-analytical modelling.

By Robinson M, Palmer S, Sculpher M, Philips Z, Ginnelly L, Bowens A, *et al.*

No. 28

Outcomes of electrically stimulated gracilis neosphincter surgery.

By Tillin T, Chambers M, Feldman R.

No. 29

The effectiveness and cost-effectiveness of pimecrolimus and tacrolimus for atopic eczema: a systematic review and economic evaluation.

By Garside R, Stein K, Castelnuovo E, Pitt M, Ashcroft D, Dimmock P, *et al.*

No. 30

Systematic review on urine albumin testing for early detection of diabetic complications.

By Newman DJ, Mattock MB, Dawnay ABS, Kerry S, McGuire A, Yaqoob M, *et al.*

No. 31

Randomised controlled trial of the cost-effectiveness of water-based therapy for lower limb osteoarthritis.

By Cochrane T, Davey RC, Matthes Edwards SM.

No. 32

Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain.

By Thomas KJ, MacPherson H, Ratcliffe J, Thorpe L, Brazier J, Campbell M, *et al.*

No. 33

Cost-effectiveness and safety of epidural steroids in the management of sciatica.

By Price C, Arden N, Coglan L, Rogers P.

No. 34

The British Rheumatoid Outcome Study Group (BROSG) randomised controlled trial to compare the effectiveness and cost-effectiveness of aggressive versus symptomatic therapy in established rheumatoid arthritis.

By Symmons D, Tricker K, Roberts C, Davies L, Dawes P, Scott DL.

No. 35

Conceptual framework and systematic review of the effects of participants' and professionals' preferences in randomised controlled trials.

By King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, *et al.*

No. 36

The clinical and cost-effectiveness of implantable cardioverter defibrillators: a systematic review.

By Bryant J, Brodin H, Loveman E, Payne E, Clegg A.

No. 37

A trial of problem-solving by community mental health nurses for anxiety, depression and life difficulties among general practice patients. The CPN-GP study.

By Kendrick T, Simons L, Mynors-Wallis L, Gray A, Lathlean J, Pickering R, *et al.*

No. 38

The causes and effects of socio-demographic exclusions from clinical trials.

By Bartlett C, Doyal L, Ebrahim S, Davey P, Bachmann M, Egger M, *et al.*

No. 39

Is hydrotherapy cost-effective? A randomised controlled trial of combined hydrotherapy programmes compared with physiotherapy land techniques in children with juvenile idiopathic arthritis.

By Epps H, Ginnelly L, Utley M, Southwood T, Gallivan S, Sculpher M, *et al.*

No. 40

A randomised controlled trial and cost-effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in people aged 65 and over. The SAFE study.

By Hobbs FDR, Fitzmaurice DA, Mant J, Murray E, Jowett S, Bryan S, *et al.*

No. 41

Displaced intracapsular hip fractures in fit, older people: a randomised comparison of reduction and fixation, bipolar hemiarthroplasty and total hip arthroplasty.

By Keating JF, Grant A, Masson M, Scott NW, Forbes JF.

No. 42

Long-term outcome of cognitive behaviour therapy clinical trials in central Scotland.

By Durham RC, Chambers JA, Power KG, Sharp DM, Macdonald RR, Major KA, *et al.*

No. 43

The effectiveness and cost-effectiveness of dual-chamber pacemakers compared with single-chamber pacemakers for bradycardia due to atrioventricular block or sick sinus syndrome: systematic review and economic evaluation.

By Castelnuovo E, Stein K, Pitt M, Garside R, Payne E.

No. 44

Newborn screening for congenital heart defects: a systematic review and cost-effectiveness analysis.

By Knowles R, Griebisch I, Dezateux C, Brown J, Bull C, Wren C.

No. 45

The clinical and cost-effectiveness of left ventricular assist devices for end-stage heart failure: a systematic review and economic evaluation.

By Clegg AJ, Scott DA, Loveman E, Colquitt J, Hutchinson J, Royle P, *et al.*

No. 46

The effectiveness of the Heidelberg Retina Tomograph and laser diagnostic glaucoma scanning system (GDx) in detecting and monitoring glaucoma.

By Kwartz AJ, Henson DB, Harper RA, Spencer AF, McLeod D.

No. 47

Clinical and cost-effectiveness of autologous chondrocyte implantation for cartilage defects in knee joints: systematic review and economic evaluation.

By Clar C, Cummins E, McIntyre L, Thomas S, Lamb J, Bain L, *et al.*

No. 48

Systematic review of effectiveness of different treatments for childhood retinoblastoma.

By McDaid C, Hartley S, Bagnall A-M, Ritchie G, Light K, Riemsma R.

No. 49

Towards evidence-based guidelines for the prevention of venous thromboembolism: systematic reviews of mechanical methods, oral anticoagulation, dextran and regional anaesthesia as thromboprophylaxis.

By Roderick P, Ferris G, Wilson K, Halls H, Jackson D, Collins R, *et al.*

No. 50

The effectiveness and cost-effectiveness of parent training/education programmes for the treatment of conduct disorder, including oppositional defiant disorder, in children.

By Dretzke J, Frew E, Davenport C, Barlow J, Stewart-Brown S, Sandercock J, *et al.*

Volume 10, 2006

No. 1

The clinical and cost-effectiveness of donepezil, rivastigmine, galantamine and memantine for Alzheimer's disease.

By Loveman E, Green C, Kirby J, Takeda A, Picot J, Payne E, *et al.*

No. 2

FOOD: a multicentre randomised trial evaluating feeding policies in patients admitted to hospital with a recent stroke.

By Dennis M, Lewis S, Cranswick G, Forbes J.

No. 3

The clinical effectiveness and cost-effectiveness of computed tomography screening for lung cancer: systematic reviews.

By Black C, Bagust A, Boland A, Walker S, McLeod C, De Verteuil R, *et al.*

No. 4

A systematic review of the effectiveness and cost-effectiveness of neuroimaging assessments used to visualise the seizure focus in people with refractory epilepsy being considered for surgery.

By Whiting P, Gupta R, Burch J, Mujica Mota RE, Wright K, Marson A, *et al.*

No. 5

Comparison of conference abstracts and presentations with full-text articles in the health technology assessments of rapidly evolving technologies.

By Dundar Y, Dodd S, Dickson R, Walley T, Haycox A, Williamson PR.

No. 6

Systematic review and evaluation of methods of assessing urinary incontinence.

By Martin JL, Williams KS, Abrams KR, Turner DA, Sutton AJ, Chapple C, *et al.*

No. 7

The clinical effectiveness and cost-effectiveness of newer drugs for children with epilepsy. A systematic review.

By Connock M, Frew E, Evans B-W, Bryan S, Cummins C, Fry-Smith A, *et al.*

No. 8

Surveillance of Barrett's oesophagus: exploring the uncertainty through systematic review, expert workshop and economic modelling.

By Garside R, Pitt M, Somerville M, Stein K, Price A, Gilbert N.

No. 9

Topotecan, pegylated liposomal doxorubicin hydrochloride and paclitaxel for second-line or subsequent treatment of advanced ovarian cancer: a systematic review and economic evaluation.

By Main C, Bojke L, Griffin S, Norman G, Barbieri M, Mather L, *et al.*

No. 10

Evaluation of molecular techniques in prediction and diagnosis of cytomegalovirus disease in immunocompromised patients.

By Szczepura A, Westmoreland D, Vinogradova Y, Fox J, Clark M.

No. 11

Screening for thrombophilia in high-risk situations: systematic review and cost-effectiveness analysis. The Thrombosis: Risk and Economic Assessment of Thrombophilia Screening (TREATS) study.

By Wu O, Robertson L, Twaddle S, Lowe GDO, Clark P, Greaves M, *et al.*

No. 12

A series of systematic reviews to inform a decision analysis for sampling and treating infected diabetic foot ulcers.

By Nelson EA, O'Meara S, Craig D, Iglesias C, Golder S, Dalton J, *et al.*

No. 13

Randomised clinical trial, observational study and assessment of cost-effectiveness of the treatment of varicose veins (REACTIV trial).

By Michaels JA, Campbell WB, Brazier JE, MacIntyre JB, Palfreyman SJ, Ratcliffe J, *et al.*

No. 14

The cost-effectiveness of screening for oral cancer in primary care.

By Speight PM, Palmer S, Moles DR, Downer MC, Smith DH, Henriksson M, *et al.*

No. 15

Measurement of the clinical and cost-effectiveness of non-invasive diagnostic testing strategies for deep vein thrombosis.

By Goodacre S, Sampson F, Stevenson M, Wailoo A, Sutton A, Thomas S, *et al.*

No. 16

Systematic review of the effectiveness and cost-effectiveness of HealOzone® for the treatment of occlusal pit/fissure caries and root caries.

By Brazzelli M, McKenzie L, Fielding S, Fraser C, Clarkson J, Kilonzo M, *et al.*

No. 17

Randomised controlled trials of conventional antipsychotic versus new atypical drugs, and new atypical drugs versus clozapine, in people with schizophrenia responding poorly to, or intolerant of, current drug treatment.

By Lewis SW, Davies L, Jones PB, Barnes TRE, Murray RM, Kerwin R, *et al.*

No. 18

Diagnostic tests and algorithms used in the investigation of haematuria: systematic reviews and economic evaluation.

By Rodgers M, Nixon J, Hempel S, Aho T, Kelly J, Neal D, *et al.*

No. 19

Cognitive behavioural therapy in addition to antispasmodic therapy for irritable bowel syndrome in primary care: randomised controlled trial.

By Kennedy TM, Chalder T, McCrone P, Darnley S, Knapp M, Jones RH, *et al.*

No. 20

A systematic review of the clinical effectiveness and cost-effectiveness of enzyme replacement therapies for Fabry's disease and mucopolysaccharidosis type 1.

By Connock M, Juarez-Garcia A, Frew E, Mans A, Dretzke J, Fry-Smith A, *et al.*

No. 21

Health benefits of antiviral therapy for mild chronic hepatitis C: randomised controlled trial and economic evaluation.

By Wright M, Grieve R, Roberts J, Main J, Thomas HC, on behalf of the UK Mild Hepatitis C Trial Investigators.

No. 22

Pressure relieving support surfaces: a randomised evaluation.

By Nixon J, Nelson EA, Cranny G, Iglesias CP, Hawkins K, Cullum NA, *et al.*

No. 23

A systematic review and economic model of the effectiveness and cost-effectiveness of methylphenidate, dexamfetamine and atomoxetine for the treatment of attention deficit hyperactivity disorder in children and adolescents.

By King S, Griffin S, Hodges Z, Weatherly H, Asseburg C, Richardson G, *et al.*

No. 24

The clinical effectiveness and cost-effectiveness of enzyme replacement therapy for Gaucher's disease: a systematic review.

By Connock M, Burls A, Frew E, Fry-Smith A, Juarez-Garcia A, McCabe C, *et al.*

No. 25

Effectiveness and cost-effectiveness of salicylic acid and cryotherapy for cutaneous warts. An economic decision model.

By Thomas KS, Keogh-Brown MR, Chalmers JR, Fordham RJ, Holland RC, Armstrong SJ, *et al.*

No. 26

A systematic literature review of the effectiveness of non-pharmacological interventions to prevent wandering in dementia and evaluation of the ethical implications and acceptability of their use.

By Robinson L, Hutchings D, Corner L, Beyer F, Dickinson H, Vanoli A, *et al.*

No. 27

A review of the evidence on the effects and costs of implantable cardioverter defibrillator therapy in different patient groups, and modelling of cost-effectiveness and cost-utility for these groups in a UK context.

By Buxton M, Caine N, Chase D, Connelly D, Grace A, Jackson C, *et al.*

No. 28

Adefovir dipivoxil and pegylated interferon alfa-2a for the treatment of chronic hepatitis B: a systematic review and economic evaluation.

By Shepherd J, Jones J, Takeda A, Davidson P, Price A.

No. 29

An evaluation of the clinical and cost-effectiveness of pulmonary artery catheters in patient management in intensive care: a systematic review and a randomised controlled trial.

By Harvey S, Stevens K, Harrison D, Young D, Brampton W, McCabe C, *et al.*

No. 30

Accurate, practical and cost-effective assessment of carotid stenosis in the UK.

By Wardlaw JM, Chappell FM, Stevenson M, De Nigris E, Thomas S, Gillard J, *et al.*

No. 31

Etanercept and infliximab for the treatment of psoriatic arthritis: a systematic review and economic evaluation.

By Woolacott N, Bravo Vergel Y, Hawkins N, Kainth A, Khadjesari Z, Misso K, *et al.*

No. 32

The cost-effectiveness of testing for hepatitis C in former injecting drug users.

By Castelnovo E, Thompson-Coon J, Pitt M, Cramp M, Siebert U, Price A, *et al.*

No. 33

Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation.

By Kaltenthaler E, Brazier J, De Nigris E, Tumor I, Ferriter M, Beverley C, *et al.*

No. 34

Cost-effectiveness of using prognostic information to select women with breast cancer for adjuvant systemic therapy.

By Williams C, Brunskill S, Altman D, Briggs A, Campbell H, Clarke M, *et al.*

No. 35

Psychological therapies including dialectical behaviour therapy for borderline personality disorder: a systematic review and preliminary economic evaluation.

By Brazier J, Tumor I, Holmes M, Ferriter M, Parry G, Dent-Brown K, *et al.*

No. 36

Clinical effectiveness and cost-effectiveness of tests for the diagnosis and investigation of urinary tract infection in children: a systematic review and economic model.

By Whiting P, Westwood M, Bojke L, Palmer S, Richardson G, Cooper J, *et al.*

No. 37

Cognitive behavioural therapy in chronic fatigue syndrome: a randomised controlled trial of an outpatient group programme.

By O'Dowd H, Gladwell P, Rogers CA, Hollinghurst S, Gregory A.

No. 38

A comparison of the cost-effectiveness of five strategies for the prevention of nonsteroidal anti-inflammatory drug-induced gastrointestinal toxicity: a systematic review with economic modelling.

By Brown TJ, Hooper L, Elliott RA, Payne K, Webb R, Roberts C, *et al.*

No. 39

The effectiveness and cost-effectiveness of computed tomography screening for coronary artery disease: systematic review.

By Waugh N, Black C, Walker S, McIntyre L, Cummins E, Hillis G.

No. 40

What are the clinical outcome and cost-effectiveness of endoscopy undertaken by nurses when compared with doctors? A Multi-Institution Nurse Endoscopy Trial (MINuET).

By Williams J, Russell I, Durai D, Cheung W-Y, Farrin A, Bloor K, *et al.*

No. 41

The clinical and cost-effectiveness of oxaliplatin and capecitabine for the adjuvant treatment of colon cancer: systematic review and economic evaluation.

By Pandor A, Eggington S, Paisley S, Tappenden P, Sutcliffe P.

No. 42

A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness.

By Chen Y-F, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, *et al.*

No. 43

Telemedicine in dermatology: a randomised controlled trial.

By Bowns IR, Collins K, Walters SJ, McDonagh AJG.

No. 44

Cost-effectiveness of cell salvage and alternative methods of minimising perioperative allogeneic blood transfusion: a systematic review and economic model.

By Davies L, Brown TJ, Haynes S, Payne K, Elliott RA, McCollum C.

No. 45

Clinical effectiveness and cost-effectiveness of laparoscopic surgery for colorectal cancer: systematic reviews and economic evaluation.

By Murray A, Lourenco T, de Verteuil R, Hernandez R, Fraser C, McKinley A, *et al.*

No. 46

Etanercept and efalizumab for the treatment of psoriasis: a systematic review.

By Woolacott N, Hawkins N, Mason A, Kainth A, Khadjesari Z, Bravo Vergel Y, *et al.*

No. 47

Systematic reviews of clinical decision tools for acute abdominal pain.

By Liu JLY, Wyatt JC, Deeks JJ, Clamp S, Keen J, Verde P, *et al.*

No. 48

Evaluation of the ventricular assist device programme in the UK.

By Sharples L, Buxton M, Caine N, Cafferty F, Demiris N, Dyer M, *et al.*

No. 49

A systematic review and economic model of the clinical and cost-effectiveness of immunosuppressive therapy for renal transplantation in children.

By Yao G, Albon E, Adi Y, Milford D, Bayliss S, Ready A, *et al.*

No. 50

Amniocentesis results: investigation of anxiety. The ARIA trial.

By Hewison J, Nixon J, Fountain J, Cocks K, Jones C, Mason G, *et al.*

Volume II, 2007

No. 1

Pemetrexed disodium for the treatment of malignant pleural mesothelioma: a systematic review and economic evaluation.

By Dundar Y, Bagust A, Dickson R, Dodd S, Green J, Haycox A, *et al.*

No. 2

A systematic review and economic model of the clinical effectiveness and cost-effectiveness of docetaxel in combination with prednisone or prednisolone for the treatment of hormone-refractory metastatic prostate cancer.

By Collins R, Fenwick E, Trowman R, Perard R, Norman G, Light K, *et al.*

No. 3

A systematic review of rapid diagnostic tests for the detection of tuberculosis infection.

By Dinnes J, Deeks J, Kunst H, Gibson A, Cummins E, Waugh N, *et al.*

No. 4

The clinical effectiveness and cost-effectiveness of strontium ranelate for the prevention of osteoporotic fragility fractures in postmenopausal women.

By Stevenson M, Davis S, Lloyd-Jones M, Beverley C.

No. 5

A systematic review of quantitative and qualitative research on the role and effectiveness of written information available to patients about individual medicines.

By Raynor DK, Blenkinsopp A, Knapp P, Grime J, Nicolson DJ, Pollock K, *et al.*

No. 6

Oral naltrexone as a treatment for relapse prevention in formerly opioid-dependent drug users: a systematic review and economic evaluation.

By Adi Y, Juarez-Garcia A, Wang D, Jowett S, Frew E, Day E, *et al.*

No. 7

Glucocorticoid-induced osteoporosis: a systematic review and cost-utility analysis.

By Kanis JA, Stevenson M, McCloskey EV, Davis S, Lloyd-Jones M.

No. 8

Epidemiological, social, diagnostic and economic evaluation of population screening for genital chlamydial infection.

By Low N, McCarthy A, Macleod J, Salisbury C, Campbell R, Roberts TE, *et al.*

No. 9

Methadone and buprenorphine for the management of opioid dependence: a systematic review and economic evaluation.

By Connock M, Juarez-Garcia A, Jowett S, Frew E, Liu Z, Taylor RJ, *et al.*

No. 10

Exercise Evaluation Randomised Trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only.

By Isaacs AJ, Critchley JA, See Tai S, Buckingham K, Westley D, Harridge SDR, *et al.*

No. 11

Interferon alfa (pegylated and non-pegylated) and ribavirin for the treatment of mild chronic hepatitis C: a systematic review and economic evaluation.

By Shepherd J, Jones J, Hartwell D, Davidson P, Price A, Waugh N.

No. 12

Systematic review and economic evaluation of bevacizumab and cetuximab for the treatment of metastatic colorectal cancer.

By Tappenden P, Jones R, Paisley S, Carroll C.

No. 13

A systematic review and economic evaluation of epoetin alfa, epoetin beta and darbepoetin alfa in anaemia associated with cancer, especially that attributable to cancer treatment.

By Wilson J, Yao GL, Raftery J, Bohlius J, Brunskill S, Sandercock J, *et al.*

No. 14

A systematic review and economic evaluation of statins for the prevention of coronary events.

By Ward S, Lloyd Jones M, Pandor A, Holmes M, Ara R, Ryan A, *et al.*

No. 15

A systematic review of the effectiveness and cost-effectiveness of different models of community-based respite care for frail older people and their carers.

By Mason A, Weatherly H, Spilsbury K, Arksey H, Golder S, Adamson J, *et al.*

No. 16

Additional therapy for young children with spastic cerebral palsy: a randomised controlled trial.

By Weindling AM, Cunningham CC, Glenn SM, Edwards RT, Reeves DJ.

No. 17

Screening for type 2 diabetes: literature review and economic modelling.

By Waugh N, Scotland G, McNamee P, Gillett M, Brennan A, Goyder E, *et al.*

No. 18

The effectiveness and cost-effectiveness of cinacalcet for secondary hyperparathyroidism in end-stage renal disease patients on dialysis: a systematic review and economic evaluation.

By Garside R, Pitt M, Anderson R, Mealing S, Roome C, Snaith A, *et al.*

No. 19

The clinical effectiveness and cost-effectiveness of gemcitabine for metastatic breast cancer: a systematic review and economic evaluation.

By Takeda AL, Jones J, Loveman E, Tan SC, Clegg AJ.

No. 20

A systematic review of duplex ultrasound, magnetic resonance angiography and computed tomography angiography for the diagnosis and assessment of symptomatic, lower limb peripheral arterial disease.

By Collins R, Cranny G, Burch J, Aguiar-Ibáñez R, Craig D, Wright K, *et al.*

No. 21

The clinical effectiveness and cost-effectiveness of treatments for children with idiopathic steroid-resistant nephrotic syndrome: a systematic review.

By Colquitt JL, Kirby J, Green C, Cooper K, Trompeter RS.

No. 22

A systematic review of the routine monitoring of growth in children of primary school age to identify growth-related conditions.

By Fayer D, Nixon J, Hartley S, Rithalia A, Butler G, Rudolf M, *et al.*

No. 23

Systematic review of the effectiveness of preventing and treating *Staphylococcus aureus* carriage in reducing peritoneal catheter-related infections.

By McCormack K, Rabindranath K, Kilonzo M, Vale L, Fraser C, McIntyre L, *et al.*

No. 24

The clinical effectiveness and cost of repetitive transcranial magnetic stimulation versus electroconvulsive therapy in severe depression: a multicentre pragmatic randomised controlled trial and economic analysis.

By McLoughlin DM, Mogg A, Eranti S, Pluck G, Purvis R, Edwards D, *et al.*

No. 25

A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment.

By Boyle J, McCartney E, Forbes J, O'Hare A.

No. 26

Hormonal therapies for early breast cancer: systematic review and economic evaluation.

By Hind D, Ward S, De Nigris E, Simpson E, Carroll C, Wyld L.

No. 27

Cardioprotection against the toxic effects of anthracyclines given to children with cancer: a systematic review.

By Bryant J, Picot J, Levitt G, Sullivan I, Baxter L, Clegg A.

No. 28

Adalimumab, etanercept and infliximab for the treatment of ankylosing spondylitis: a systematic review and economic evaluation.

By McLeod C, Bagust A, Boland A, Dagenais P, Dickson R, Dundar Y, *et al.*

No. 29

Prenatal screening and treatment strategies to prevent group B streptococcal and other bacterial infections in early infancy: cost-effectiveness and expected value of information analyses.

By Colbourn T, Asseburg C, Bojke L, Philips Z, Claxton K, Ades AE, *et al.*

No. 30

Clinical effectiveness and cost-effectiveness of bone morphogenetic proteins in the non-healing of fractures and spinal fusion: a systematic review.

By Garrison KR, Donell S, Ryder J, Shemilt I, Mugford M, Harvey I, *et al.*

No. 31

A randomised controlled trial of postoperative radiotherapy following breast-conserving surgery in a minimum-risk older population. The PRIME trial.

By Prescott RJ, Kunkler IH, Williams LJ, King CC, Jack W, van der Pol M, *et al.*

No. 32

Current practice, accuracy, effectiveness and cost-effectiveness of the school entry hearing screen.

By Bamford J, Fortnum H, Bristow K, Smith J, Vamvakas G, Davies L, *et al.*

No. 33

The clinical effectiveness and cost-effectiveness of inhaled insulin in diabetes mellitus: a systematic review and economic evaluation.

By Black C, Cummins E, Royle P, Philip S, Waugh N.

No. 34

Surveillance of cirrhosis for hepatocellular carcinoma: systematic review and economic analysis.

By Thompson Coon J, Rogers G, Hewson P, Wright D, Anderson R, Cramp M, *et al.*

No. 35

The Birmingham Rehabilitation Uptake Maximisation Study (BRUM). Homebased compared with hospital-based cardiac rehabilitation in a multi-ethnic population: cost-effectiveness and patient adherence.

By Jolly K, Taylor R, Lip GYH, Greenfield S, Raftery J, Mant J, *et al.*

No. 36

A systematic review of the clinical, public health and cost-effectiveness of rapid diagnostic tests for the detection and identification of bacterial intestinal pathogens in faeces and food.

By Abubakar I, Irvine L, Aldus CF, Wyatt GM, Fordham R, Schelenz S, *et al.*

No. 37

A randomised controlled trial examining the longer-term outcomes of standard versus new antiepileptic drugs. The SANAD trial.

By Marson AG, Appleton R, Baker GA, Chadwick DW, Doughty J, Eaton B, *et al.*

No. 38

Clinical effectiveness and cost-effectiveness of different models of managing long-term oral anti-coagulation therapy: a systematic review and economic modelling.

By Connock M, Stevens C, Fry-Smith A, Jowett S, Fitzmaurice D, Moore D, *et al.*

No. 39

A systematic review and economic model of the clinical effectiveness and cost-effectiveness of interventions for preventing relapse in people with bipolar disorder.

By Soares-Weiser K, Bravo Vergel Y, Beynon S, Dunn G, Barbieri M, Duffy S, *et al.*

No. 40

Taxanes for the adjuvant treatment of early breast cancer: systematic review and economic evaluation.

By Ward S, Simpson E, Davis S, Hind D, Rees A, Wilkinson A.

No. 41

The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation.

By Burr JM, Mowatt G, Hernández R, Siddiqui MAR, Cook J, Lourenco T, *et al.*

No. 42

Acceptability, benefit and costs of early screening for hearing disability: a study of potential screening tests and models.

By Davis A, Smith P, Ferguson M, Stephens D, Gianopoulos I.

No. 43

Contamination in trials of educational interventions.

By Keogh-Brown MR, Bachmann MO, Shepstone L, Hewitt C, Howe A, Ramsay CR, *et al.*

No. 44

Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers.

By Facey K, Bradbury I, Laking G, Payne E.

No. 45

The effectiveness and cost-effectiveness of carmustine implants and temozolomide for the treatment of newly diagnosed high-grade glioma: a systematic review and economic evaluation.

By Garside R, Pitt M, Anderson R, Rogers G, Dyer M, Mealing S, *et al.*

No. 46

Drug-eluting stents: a systematic review and economic evaluation.

By Hill RA, Boland A, Dickson R, Dünder Y, Haycox A, McLeod C, *et al.*

No. 47

The clinical effectiveness and cost-effectiveness of cardiac resynchronisation (biventricular pacing) for heart failure: systematic review and economic model.

By Fox M, Mealing S, Anderson R, Dean J, Stein K, Price A, *et al.*

No. 48

Recruitment to randomised trials: strategies for trial enrolment and participation study. The STEPS study.

By Campbell MK, Snowden C, Francis D, Elbourne D, McDonald AM, Knight R, *et al.*

No. 49

Cost-effectiveness of functional cardiac testing in the diagnosis and management of coronary artery disease: a randomised controlled trial. The CECaT trial.

By Sharples L, Hughes V, Crean A, Dyer M, Buxton M, Goldsmith K, *et al.*

No. 50

Evaluation of diagnostic tests when there is no gold standard. A review of methods.

By Rutjes AWS, Reitsma JB, Coomarasamy A, Khan KS, Bossuyt PMM.

No. 51

Systematic reviews of the clinical effectiveness and cost-effectiveness of proton pump inhibitors in acute upper gastrointestinal bleeding.

By Leontiadis GI, Sreedharan A, Dorward S, Barton P, Delaney B, Howden CW, *et al.*

No. 52

A review and critique of modelling in prioritising and designing screening programmes.

By Karnon J, Goyder E, Tappenden P, McPhie S, Towers I, Brazier J, *et al.*

No. 53

An assessment of the impact of the NHS Health Technology Assessment Programme.

By Hanney S, Buxton M, Green C, Coulson D, Raftery J.

Volume 12, 2008

No. 1

A systematic review and economic model of switching from nonglycopeptide to glycopeptide antibiotic prophylaxis for surgery.

By Cranny G, Elliott R, Weatherly H, Chambers D, Hawkins N, Myers L, *et al.*

No. 2

'Cut down to quit' with nicotine replacement therapies in smoking cessation: a systematic review of effectiveness and economic analysis.

By Wang D, Connock M, Barton P, Fry-Smith A, Aveyard P, Moore D.

No. 3

A systematic review of the effectiveness of strategies for reducing fracture risk in children with juvenile idiopathic arthritis with additional data on long-term risk of fracture and cost of disease management.

By Thornton J, Ashcroft D, O'Neill T, Elliott R, Adams J, Roberts C, *et al.*

No. 4

Does befriending by trained lay workers improve psychological well-being and quality of life for carers of people with dementia, and at what cost? A randomised controlled trial.

By Charlesworth G, Shepstone L, Wilson E, Thalany M, Mugford M, Poland F.

No. 5

A multi-centre retrospective cohort study comparing the efficacy, safety and cost-effectiveness of hysterectomy and uterine artery embolisation for the treatment of symptomatic uterine fibroids. The HOPEFUL study.

By Hirst A, Dutton S, Wu O, Briggs A, Edwards C, Waldenmaier L, *et al.*

No. 6

Methods of prediction and prevention of pre-eclampsia: systematic reviews of accuracy and effectiveness literature with economic modelling.

By Meads CA, Cnossen JS, Meher S, Juarez-Garcia A, ter Riet G, Duley L, *et al.*

No. 7

The use of economic evaluations in NHS decision-making: a review and empirical investigation.

By Williams I, McIver S, Moore D, Bryan S.

No. 8

Stapled haemorrhoidectomy (haemorrhoidopexy) for the treatment of haemorrhoids: a systematic review and economic evaluation.

By Burch J, Epstein D, Baba-Akbari A, Weatherly H, Fox D, Golder S, *et al.*

No. 9

The clinical effectiveness of diabetes education models for Type 2 diabetes: a systematic review.

By Loveman E, Frampton GK, Clegg AJ.

No. 10

Payment to healthcare professionals for patient recruitment to trials: systematic review and qualitative study.

By Raftery J, Bryant J, Powell J, Kerr C, Hawker S.

No. 11

Cyclooxygenase-2 selective non-steroidal anti-inflammatory drugs (etodolac, meloxicam, celecoxib, rofecoxib, etoricoxib, valdecoxib and lumiracoxib) for osteoarthritis and rheumatoid arthritis: a systematic review and economic evaluation.

By Chen Y-F, Jobanputra P, Barton P, Bryan S, Fry-Smith A, Harris G, *et al.*

No. 12

The clinical effectiveness and cost-effectiveness of central venous catheters treated with anti-infective agents in preventing bloodstream infections: a systematic review and economic evaluation.

By Hockenhull JC, Dwan K, Boland A, Smith G, Bagust A, Dundar Y, *et al.*

No. 13

Stepped treatment of older adults on laxatives. The STOOL trial.

By Mihaylov S, Stark C, McColl E, Steen N, Vanoli A, Rubin G, *et al.*

No. 14

A randomised controlled trial of cognitive behaviour therapy in adolescents with major depression treated by selective serotonin reuptake inhibitors. The ADAPT trial.

By Goodyer IM, Dubicka B, Wilkinson P, Kelvin R, Roberts C, Byford S, *et al.*

No. 15

The use of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer: systematic review and economic evaluation.

By Hind D, Tappenden P, Tumor I, Eggington E, Sutcliffe P, Ryan A.

No. 16

Ranibizumab and pegaptanib for the treatment of age-related macular degeneration: a systematic review and economic evaluation.

By Colquitt JL, Jones J, Tan SC, Takeda A, Clegg AJ, Price A.

No. 17

Systematic review of the clinical effectiveness and cost-effectiveness of 64-slice or higher computed tomography angiography as an alternative to invasive coronary angiography in the investigation of coronary artery disease.

By Mowatt G, Cummins E, Waugh N, Walker S, Cook J, Jia X, *et al.*

No. 18

Structural neuroimaging in psychosis: a systematic review and economic evaluation.

By Albon E, Tsourapas A, Frew E, Davenport C, Oyeboode F, Bayliss S, *et al.*

No. 19

Systematic review and economic analysis of the comparative effectiveness of different inhaled corticosteroids and their usage with long-acting beta₂ agonists for the treatment of chronic asthma in adults and children aged 12 years and over.

By Shepherd J, Rogers G, Anderson R, Main C, Thompson-Coon J, Hartwell D, *et al.*

No. 20

Systematic review and economic analysis of the comparative effectiveness of different inhaled corticosteroids and their usage with long-acting beta₂ agonists for the treatment of chronic asthma in children under the age of 12 years.

By Main C, Shepherd J, Anderson R, Rogers G, Thompson-Coon J, Liu Z, *et al.*

No. 21

Ezetimibe for the treatment of hypercholesterolaemia: a systematic review and economic evaluation.

By Ara R, Tumor I, Pandor A, Duenas A, Williams R, Wilkinson A, *et al.*

No. 22

Topical or oral ibuprofen for chronic knee pain in older people. The TOIB study.

By Underwood M, Ashby D, Carnes D, Castelnuovo E, Cross P, Harding G, *et al.*

No. 23

A prospective randomised comparison of minor surgery in primary and secondary care. The MiSTIC trial.

By George S, Pockney P, Primrose J, Smith H, Little P, Kinley H, *et al.*

No. 24

A review and critical appraisal of measures of therapist-patient interactions in mental health settings.

By Cahill J, Barkham M, Hardy G, Gilbody S, Richards D, Bower P, *et al.*

No. 25

The clinical effectiveness and cost-effectiveness of screening programmes for amblyopia and strabismus in children up to the age of 4–5 years: a systematic review and economic evaluation.

By Carlton J, Karnon J, Czoski-Murray C, Smith KJ, Marr J.

No. 26

A systematic review of the clinical effectiveness and cost-effectiveness and economic modelling of minimal incision total hip replacement approaches in the management of arthritic disease of the hip.

By de Verteuil R, Imamura M, Zhu S, Glazener C, Fraser C, Munro N, *et al.*

No. 27

A preliminary model-based assessment of the cost-utility of a screening programme for early age-related macular degeneration.

By Karnon J, Czoski-Murray C, Smith K, Brand C, Chakravarthy U, Davis S, *et al.*

No. 28

Intravenous magnesium sulphate and sotalol for prevention of atrial fibrillation after coronary artery bypass surgery: a systematic review and economic evaluation.

By Shepherd J, Jones J, Frampton GK, Tanajewski L, Turner D, Price A.

No. 29

Absorbent products for urinary/faecal incontinence: a comparative evaluation of key product categories.

By Fader M, Cottenden A, Getliffe K, Gage H, Clarke-O'Neill S, Jamieson K, *et al.*

No. 30

A systematic review of repetitive functional task practice with modelling of resource use, costs and effectiveness.

By French B, Leathley M, Sutton C, McAdam J, Thomas L, Forster A, *et al.*

No. 31

The effectiveness and cost-effectiveness of minimal access surgery amongst people with gastro-oesophageal reflux disease – a UK collaborative study. The REFLUX trial.

By Grant A, Wileman S, Ramsay C, Bojke L, Epstein D, Sculpher M, *et al.*

No. 32

Time to full publication of studies of anti-cancer medicines for breast cancer and the potential for publication bias: a short systematic review.

By Takeda A, Loveman E, Harris P, Hartwell D, Welch K.

No. 33

Performance of screening tests for child physical abuse in accident and emergency departments.

By Woodman J, Pitt M, Wentz R, Taylor B, Hodes D, Gilbert RE.

No. 34

Curative catheter ablation in atrial fibrillation and typical atrial flutter: systematic review and economic evaluation.

By Rodgers M, McKenna C, Palmer S, Chambers D, Van Hout S, Golder S, *et al.*

No. 35

Systematic review and economic modelling of effectiveness and cost utility of surgical treatments for men with benign prostatic enlargement.

By Lourenco T, Armstrong N, N'Dow J, Nabi G, Deverill M, Pickard R, *et al.*

No. 36

Immunoprophylaxis against respiratory syncytial virus (RSV) with palivizumab in children: a systematic review and economic evaluation.

By Wang D, Cummins C, Bayliss S, Sandercock J, Burls A.

Volume 13, 2009**No. 1**

Deferasirox for the treatment of iron overload associated with regular blood transfusions (transfusional haemosiderosis) in patients suffering with chronic anaemia: a systematic review and economic evaluation.

By McLeod C, Fleeman N, Kirkham J, Bagust A, Boland A, Chu P, *et al.*

No. 2

Thrombophilia testing in people with venous thromboembolism: systematic review and cost-effectiveness analysis.

By Simpson EL, Stevenson MD, Rawdin A, Papaioannou D.

No. 3

Surgical procedures and non-surgical devices for the management of non-apnoeic snoring: a systematic review of clinical effects and associated treatment costs.

By Main C, Liu Z, Welch K, Weiner G, Quentin Jones S, Stein K.

No. 4

Continuous positive airway pressure devices for the treatment of obstructive sleep apnoea-hypopnoea syndrome: a systematic review and economic analysis.

By McDaid C, Griffin S, Weatherly H, Durée K, van der Burgt M, van Hout S, Akers J, *et al.*

No. 5

Use of classical and novel biomarkers as prognostic risk factors for localised prostate cancer: a systematic review.

By Sutcliffe P, Hummel S, Simpson E, Young T, Rees A, Wilkinson A, *et al.*

No. 6

The harmful health effects of recreational ecstasy: a systematic review of observational evidence.

By Rogers G, Elston J, Garside R, Roome C, Taylor R, Younger P, *et al.*

No. 7

Systematic review of the clinical effectiveness and cost-effectiveness of oesophageal Doppler monitoring in critically ill and high-risk surgical patients.

By Mowatt G, Houston G, Hernández R, de Verteuil R, Fraser C, Cuthbertson B, *et al.*

No. 8

The use of surrogate outcomes in model-based cost-effectiveness analyses: a survey of UK Health Technology Assessment reports.

By Taylor RS, Elston J.

No. 9

Controlling Hypertension and Hypotension Immediately Post Stroke (CHHIPS) – a randomised controlled trial.

By Potter J, Mistri A, Brodie F, Chernova J, Wilson E, Jagger C, *et al.*

No. 10

Routine antenatal anti-D prophylaxis for RhD-negative women: a systematic review and economic evaluation.

By Pilgrim H, Lloyd-Jones M, Rees A.

No. 11

Amantadine, oseltamivir and zanamivir for the prophylaxis of influenza (including a review of existing guidance no. 67): a systematic review and economic evaluation.

By Tappenden P, Jackson R, Cooper K, Rees A, Simpson E, Read R, *et al.*

No. 12

Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods.

By Hobart J, Cano S.

No. 13

Treatment of severe ankle sprain: a pragmatic randomised controlled trial comparing the clinical effectiveness and cost-effectiveness of three types of mechanical ankle support with tubular bandage. The CAST trial.

By Cooke MW, Marsh JL, Clark M, Nakash R, Jarvis RM, Hutton JL, *et al.*, on behalf of the CAST trial group.

No. 14

Non-occupational postexposure prophylaxis for HIV: a systematic review.

By Bryant J, Baxter L, Hird S.

No. 15

Blood glucose self-monitoring in type 2 diabetes: a randomised controlled trial.

By Farmer AJ, Wade AN, French DP, Simon J, Yudkin P, Gray A, *et al.*

No. 16

How far does screening women for domestic (partner) violence in different health-care settings meet criteria for a screening programme? Systematic reviews of nine UK National Screening Committee criteria.

By Feder G, Ramsay J, Dunne D, Rose M, Arsene C, Norman R, *et al.*

No. 17

Spinal cord stimulation for chronic pain of neuropathic or ischaemic origin: systematic review and economic evaluation.

By Simpson, EL, Duenas A, Holmes MW, Papaioannou D, Chilcott J.

No. 18

The role of magnetic resonance imaging in the identification of suspected acoustic neuroma: a systematic review of clinical and cost-effectiveness and natural history.

By Fortnum H, O'Neill C, Taylor R, Lenthall R, Nikolopoulos T, Lightfoot G, *et al.*

No. 19

Dipsticks and diagnostic algorithms in urinary tract infection: development and validation, randomised trial, economic analysis, observational cohort and qualitative study.

By Little P, Turner S, Rumsby K, Warner G, Moore M, Lowes JA, *et al.*

No. 20

Systematic review of respite care in the frail elderly.

By Shaw C, McNamara R, Abrams K, Cannings-John R, Hood K, Longo M, *et al.*

No. 21

Neuroleptics in the treatment of aggressive challenging behaviour for people with intellectual disabilities: a randomised controlled trial (NACHBID).

By Tyrer P, Oliver-Africano P, Romeo R, Knapp M, Dickens S, Bouras N, *et al.*

No. 22

Randomised controlled trial to determine the clinical effectiveness and cost-effectiveness of selective serotonin reuptake inhibitors plus supportive care, versus supportive care alone, for mild to moderate depression with somatic symptoms in primary care: the THREAD (THREShold for AntiDepressant response) study.

By Kendrick T, Chatwin J, Dowrick C, Tylee A, Morriss R, Peveler R, *et al.*

No. 23

Diagnostic strategies using DNA testing for hereditary haemochromatosis in at-risk populations: a systematic review and economic evaluation.

By Bryant J, Cooper K, Picot J, Clegg A, Roderick P, Rosenberg W, *et al.*

No. 24

Enhanced external counterpulsation for the treatment of stable angina and heart failure: a systematic review and economic analysis.

By McKenna C, McDaid C, Suekarran S, Hawkins N, Claxton K, Light K, *et al.*

No. 25

Development of a decision support tool for primary care management of patients with abnormal liver function tests without clinically apparent liver disease: a record-linkage population cohort study and decision analysis (ALFIE).

By Donnan PT, McLernon D, Dillon JF, Ryder S, Roderick P, Sullivan F, *et al.*

No. 26

A systematic review of presumed consent systems for deceased organ donation.

By Rithalia A, McDaid C, Suekarran S, Norman G, Myers L, Sowden A.

No. 27

Paracetamol and ibuprofen for the treatment of fever in children: the PITCH randomised controlled trial.

By Hay AD, Redmond NM, Costelloe C, Montgomery AA, Fletcher M, Hollinghurst S, *et al.*

No. 28

A randomised controlled trial to compare minimally invasive glucose monitoring devices with conventional monitoring in the management of insulin-treated diabetes mellitus (MITRE).

By Newman SP, Cooke D, Casbard A, Walker S, Meredith S, Nunn A, *et al.*

No. 29

Sensitivity analysis in economic evaluation: an audit of NICE current practice and a review of its use and value in decision-making.

By Andronis L, Barton P, Bryan S.

Suppl. 1

Trastuzumab for the treatment of primary breast cancer in HER2-positive women: a single technology appraisal.

By Ward S, Pilgrim H, Hind D.

Docetaxel for the adjuvant treatment of early node-positive breast cancer: a single technology appraisal.

By Chilcott J, Lloyd Jones M, Wilkinson A.

The use of paclitaxel in the management of early stage breast cancer.

By Griffin S, Dunn G, Palmer S, Macfarlane K, Brent S, Dyker A, *et al.*

Rituximab for the first-line treatment of stage III/IV follicular non-Hodgkin's lymphoma.

By Dundar Y, Bagust A, Hounsome J, McLeod C, Boland A, Davis H, *et al.*

Bortezomib for the treatment of multiple myeloma patients.

By Green C, Bryant J, Takeda A, Cooper K, Clegg A, Smith A, *et al.*

Fludarabine phosphate for the first-line treatment of chronic lymphocytic leukaemia.

By Walker S, Palmer S, Erhorn S, Brent S, Dyker A, Ferrie L, *et al.*

Erlotinib for the treatment of relapsed non-small cell lung cancer.

By McLeod C, Bagust A, Boland A, Hockenhull J, Dundar Y, Proudlove C, *et al.*

Cetuximab plus radiotherapy for the treatment of locally advanced squamous cell carcinoma of the head and neck.

By Griffin S, Walker S, Sculpher M, White S, Erhorn S, Brent S, *et al.*

Infliximab for the treatment of adults with psoriasis.

By Loveman E, Turner D, Hartwell D, Cooper K, Clegg A.

No. 30

Psychological interventions for postnatal depression: cluster randomised trial and economic evaluation. The PoNDER trial.

By Morrell CJ, Warner R, Slade P, Dixon S, Walters S, Paley G, *et al.*

No. 31

The effect of different treatment durations of clopidogrel in patients with non-ST-segment elevation acute coronary syndromes: a systematic review and value of information analysis.

By Rogowski R, Burch J, Palmer S, Craigs C, Golder S, Woolacott N.

No. 32

Systematic review and individual patient data meta-analysis of diagnosis of heart failure, with modelling of implications of different diagnostic strategies in primary care.

By Mant J, Doust J, Roalfe A, Barton P, Cowie MR, Glasziou P, *et al.*

No. 33

A multicentre randomised controlled trial of the use of continuous positive airway pressure and non-invasive positive pressure ventilation in the early treatment of patients presenting to the emergency department with severe acute cardiogenic pulmonary oedema: the 3CPO trial.

By Gray AJ, Goodacre S, Newby DE, Masson MA, Sampson F, Dixon S, *et al.*, on behalf of the 3CPO study investigators.

No. 34

Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation.

By Ara R, Pandor A, Stevens J, Rees A, Rafia R.

No. 35

Adefovir dipivoxil and pegylated interferon alpha for the treatment of chronic hepatitis B: an updated systematic review and economic evaluation.

By Jones J, Shepherd J, Baxter L, Gospodarevskaya E, Hartwell D, Harris P, *et al.*

No. 36

Methods to identify postnatal depression in primary care: an integrated evidence synthesis and value of information analysis.

By Hewitt CE, Gilbody SM, Brealey S, Paulden M, Palmer S, Mann R, *et al.*

No. 37

A double-blind randomised placebo-controlled trial of topical intranasal corticosteroids in 4- to 11-year-old children with persistent bilateral otitis media with effusion in primary care.

By Williamson I, Bengt S, Barton S, Petrou S, Letley L, Fasey N, *et al.*

No. 38

The effectiveness and cost-effectiveness of methods of storing donated kidneys from deceased donors: a systematic review and economic model.

By Bond M, Pitt M, Akoh J, Moxham T, Hoyle M, Anderson R.

No. 39

Rehabilitation of older patients: day hospital compared with rehabilitation at home. A randomised controlled trial.

By Parker SG, Oliver P, Pennington M, Bond J, Jagger C, Enderby PM, *et al.*

No. 40

Breastfeeding promotion for infants in neonatal units: a systematic review and economic analysis

By Renfrew MJ, Craig D, Dyson L, McCormick F, Rice S, King SE, *et al.*

No. 41

The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation.

By Picot J, Jones J, Colquitt JL, Gospodarevskaya E, Loveman E, Baxter L, *et al.*

No. 42

Rapid testing for group B streptococcus during labour: a test accuracy study with evaluation of acceptability and cost-effectiveness.

By Daniels J, Gray J, Pattison H, Roberts T, Edwards E, Milner P, *et al.*

No. 43

Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling.

By Honest H, Forbes CA, Durée KH, Norman G, Duffy SB, Tsourapas A, *et al.*

No. 44

The effectiveness and cost-effectiveness of cochlear implants for severe to profound deafness in children and adults: a systematic review and economic model.

By Bond M, Mealing S, Anderson R, Elston J, Weiner G, Taylor RS, *et al.*

Suppl. 2

Gemcitabine for the treatment of metastatic breast cancer.

By Jones J, Takeda A, Tan SC, Cooper K, Loveman E, Clegg A.

Varenicline in the management of smoking cessation: a single technology appraisal.

By Hind D, Tappenden P, Peters J, Kenjegalieva K.

Alteplase for the treatment of acute ischaemic stroke: a single technology appraisal.

By Lloyd Jones M, Holmes M.

Rituximab for the treatment of rheumatoid arthritis.

By Bagust A, Boland A, Hockenhull J, Fleeman N, Greenhalgh J, Dundar Y, *et al.*

Omalizumab for the treatment of severe persistent allergic asthma.

By Jones J, Shepherd J, Hartwell D, Harris P, Cooper K, Takeda A, *et al.*

Rituximab for the treatment of relapsed or refractory stage III or IV follicular non-Hodgkin's lymphoma.

By Boland A, Bagust A, Hockenhull J, Davis H, Chu P, Dickson R.

Adalimumab for the treatment of psoriasis.

By Turner D, Picot J, Cooper K, Loveman E.

Dabigatran etexilate for the prevention of venous thromboembolism in patients undergoing elective hip and knee surgery: a single technology appraisal.

By Holmes M, C Carroll C, Papaioannou D.

Romiplostim for the treatment of chronic immune or idiopathic thrombocytopenic purpura: a single technology appraisal.

By Mowatt G, Boachie C, Crowther M, Fraser C, Hernández R, Jia X, *et al.*

Sunitinib for the treatment of gastrointestinal stromal tumours: a critique of the submission from Pfizer.

By Bond M, Hoyle M, Moxham T, Napier M, Anderson R.

No. 45

Vitamin K to prevent fractures in older women: systematic review and economic evaluation.

By Stevenson M, Lloyd-Jones M, Papaioannou D.

No. 46

The effects of biofeedback for the treatment of essential hypertension: a systematic review.

By Greenhalgh J, Dickson R, Dundar Y.

No. 47

A randomised controlled trial of the use of aciclovir and/or prednisolone for the early treatment of Bell's palsy: the BELLS study.

By Sullivan FM, Swan IRC, Donnan PT, Morrison JM, Smith BH, McKinstry B, *et al.*

Suppl. 3

Lapatinib for the treatment of HER2-overexpressing breast cancer.

By Jones J, Takeda A, Picot J, von Keyserlingk C, Clegg A.

Infliximab for the treatment of ulcerative colitis.

By Hyde C, Bryan S, Juarez-Garcia A, Andronis L, Fry-Smith A.

Rimonabant for the treatment of overweight and obese people.

By Burch J, McKenna C, Palmer S, Norman G, Glanville J, Sculpher M, *et al.*

Telbivudine for the treatment of chronic hepatitis B infection.

By Hartwell D, Jones J, Harris P, Cooper K.

Entecavir for the treatment of chronic hepatitis B infection.

By Shepherd J, Gospodarevskaya E, Frampton G, Cooper K.

Febuxostat for the treatment of hyperuricaemia in people with gout: a single technology appraisal.

By Stevenson M, Pandor A.

Rivaroxaban for the prevention of venous thromboembolism: a single technology appraisal.

By Stevenson M, Scope A, Holmes M, Rees A, Kaltenthaler E.

Cetuximab for the treatment of recurrent and/or metastatic squamous cell carcinoma of the head and neck.

By Greenhalgh J, Bagust A, Boland A, Fleeman N, McLeod C, Dundar Y, *et al.*

Mifamurtide for the treatment of osteosarcoma: a single technology appraisal.

By Pandor A, Fitzgerald P, Stevenson M, Papaioannou D.

Ustekinumab for the treatment of moderate to severe psoriasis.

By Gospodarevskaya E, Picot J, Cooper K, Loveman E, Takeda A.

No. 48

Endovascular stents for abdominal aortic aneurysms: a systematic review and economic model.

By Chambers D, Epstein D, Walker S, Fayter D, Paton F, Wright K, *et al.*

No. 49

Clinical and cost-effectiveness of epoprostenol, iloprost, bosentan, sitaxentan and sildenafil for pulmonary arterial hypertension within their licensed indications: a systematic review and economic evaluation.

By Chen Y-F, Jowett S, Barton P, Malottki K, Hyde C, Gibbs JSR, *et al.*

No. 50

Cessation of attention deficit hyperactivity disorder drugs in the young (CADDY) – a pharmacoepidemiological and qualitative study.

By Wong ICK, Asherson P, Bilbow A, Clifford S, Coghil D, R DeSoysa R, *et al.*

No. 51

ARTISTIC: a randomised trial of human papillomavirus (HPV) testing in primary cervical screening.

By Kitchener HC, Almonte M, Gilham C, Dowie R, Stoykova B, Sargent A, *et al.*

No. 52

The clinical effectiveness of glucosamine and chondroitin supplements in slowing or arresting progression of osteoarthritis of the knee: a systematic review and economic evaluation.

By Black C, Clar C, Henderson R, MacEachern C, McNamee P, Quayyum Z, *et al.*

No. 53

Randomised preference trial of medical versus surgical termination of pregnancy less than 14 weeks' gestation (TOPS).

By Robson SC, Kelly T, Howel D, Deverill M, Hewison J, Lie MLS, *et al.*

No. 54

Randomised controlled trial of the use of three dressing preparations in the management of chronic ulceration of the foot in diabetes.

By Jeffcoate WJ, Price PE, Phillips CJ, Game FL, Mudge E, Davies S, *et al.*

No. 55

VenUS II: a randomised controlled trial of larval therapy in the management of leg ulcers.

By Dumville JC, Worthy G, Soares MO, Bland JM, Cullum N, Dowson C, *et al.*

No. 56

A prospective randomised controlled trial and economic modelling of antimicrobial silver dressings versus non-adherent control dressings for venous leg ulcers: the VULCAN trial

By Michaels JA, Campbell WB, King BM, MacIntyre J, Palfreyman SJ, Shackley P, *et al.*

No. 57

Communication of carrier status information following universal newborn screening for sickle cell disorders and cystic fibrosis: qualitative study of experience and practice.

By Kai J, Ulph F, Cullinan T, Qureshi N.

No. 58

Antiviral drugs for the treatment of influenza: a systematic review and economic evaluation.

By Burch J, Paulden M, Conti S, Stock C, Corbett M, Welton NJ, *et al.*

No. 59

Development of a toolkit and glossary to aid in the adaptation of health technology assessment (HTA) reports for use in different contexts.

By Chase D, Rosten C, Turner S, Hicks N, Milne R.

No. 60

Colour vision testing for diabetic retinopathy: a systematic review of diagnostic accuracy and economic evaluation.

By Rodgers M, Hodges R, Hawkins J, Hollingworth W, Duffy S, McKibbin M, *et al.*

No. 61

Systematic review of the effectiveness and cost-effectiveness of weight management schemes for the under fives: a short report.

By Bond M, Wyatt K, Lloyd J, Welch K, Taylor R.

No. 62

Are adverse effects incorporated in economic models? An initial review of current practice.

By Craig D, McDaid C, Fonseca T, Stock C, Duffy S, Woolacott N.

Volume 14, 2010

No. 1

Multicentre randomised controlled trial examining the cost-effectiveness of contrast-enhanced high field magnetic resonance imaging in women with primary breast cancer scheduled for wide local excision (COMICE).

By Turnbull LW, Brown SR, Olivier C, Harvey I, Brown J, Drew P, *et al.*

No. 2

Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation.

By Thompson Coon J, Hoyle M, Green C, Liu Z, Welch K, Moxham T, *et al.*

No. 3

The clinical effectiveness and cost-effectiveness of testing for cytochrome P450 polymorphisms in patients with schizophrenia treated with antipsychotics: a systematic review and economic evaluation.

By Fleeman N, McLeod C, Bagust A, Beale S, Boland A, Dundar Y, *et al.*

No. 4

Systematic review of the clinical effectiveness and cost-effectiveness of photodynamic diagnosis and urine biomarkers (FISH, ImmunoCyt, NMP22) and cytology for the detection and follow-up of bladder cancer.

By Mowatt G, Zhu S, Kilonzo M, Boachie C, Fraser C, Griffiths TRL, *et al.*

No. 5

Effectiveness and cost-effectiveness of arthroscopic lavage in the treatment of osteoarthritis of the knee: a mixed methods study of the feasibility of conducting a surgical placebo-controlled trial (the KORAL study).

By Campbell MK, Skea ZC, Sutherland AG, Cuthbertson BH, Entwistle VA, McDonald AM, *et al.*

No. 6

A randomised 2 × 2 trial of community versus hospital pulmonary rehabilitation for chronic obstructive pulmonary disease followed by telephone or conventional follow-up.

By Waterhouse JC, Walters SJ, Oluboyede Y, Lawson RA.

No. 7

The effectiveness and cost-effectiveness of behavioural interventions for the prevention of sexually transmitted infections in young people aged 13–19: a systematic review and economic evaluation.

By Shepherd J, Kavanagh J, Picot J, Cooper K, Harden A, Barnett-Page E, *et al.*

No. 8

Dissemination and publication of research findings: an updated review of related biases.

By Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, *et al.*

No. 9

The effectiveness and cost-effectiveness of biomarkers for the prioritisation of patients awaiting coronary revascularisation: a systematic review and decision model.

By Hemingway H, Henriksson M, Chen R, Damant J, Fitzpatrick N, Abrams K, *et al.*

No. 10

Comparison of case note review methods for evaluating quality and safety in health care.

By Hutchinson A, Coster JE, Cooper KL, McIntosh A, Walters SJ, Bath PA, *et al.*

No. 11

Clinical effectiveness and cost-effectiveness of continuous subcutaneous insulin infusion for diabetes: systematic review and economic evaluation.

By Cummins E, Royle P, Snaith A, Greene A, Robertson L, McIntyre L, *et al.*

No. 12

Self-monitoring of blood glucose in type 2 diabetes: systematic review.

By Clar C, Barnard K, Cummins E, Royle P, Waugh N.

No. 13

North of England and Scotland Study of Tonsillectomy and Adeno-tonsillectomy in Children (NESSTAC): a pragmatic randomised controlled trial with a parallel non-randomised preference study.

By Lock C, Wilson J, Steen N, Eccles M, Mason H, Carrie S, *et al.*

No. 14

Multicentre randomised controlled trial of the clinical and cost-effectiveness of a bypass-surgery-first versus a balloon-angioplasty-first revascularisation strategy for severe limb ischaemia due to infrainguinal disease. The Bypass versus Angioplasty in Severe Ischaemia of the Leg (BASIL) trial.

By Bradbury AW, Adam DJ, Bell J, Forbes JF, Fowkes FGR, Gillespie I, *et al.*

No. 15

A randomised controlled multicentre trial of treatments for adolescent anorexia nervosa including assessment of cost-effectiveness and patient acceptability – the TOuCAN trial.

By Gowers SG, Clark AF, Roberts C, Byford S, Barrett B, Griffiths A, *et al.*

No. 16

Randomised controlled trials for policy interventions: a review of reviews and meta-regression.

By Oliver S, Bagnall AM, Thomas J, Shepherd J, Sowden A, White I, *et al.*

No. 17

Paracetamol and selective and non-selective non-steroidal anti-inflammatory drugs (NSAIDs) for the reduction of morphine-related side effects after major surgery: a systematic review.

By McDaid C, Maund E, Rice S, Wright K, Jenkins B, Woolacott N.

No. 18

A systematic review of outcome measures used in forensic mental health research with consensus panel opinion.

By Fitzpatrick R, Chambers J, Burns T, Doll H, Fazel S, Jenkinson C, *et al.*

No. 19

The clinical effectiveness and cost-effectiveness of topotecan for small cell lung cancer: a systematic review and economic evaluation.

By Loveman E, Jones J, Hartwell D, Bird A, Harris P, Welch K, *et al.*

No. 20

Antenatal screening for haemoglobinopathies in primary care: a cohort study and cluster randomised trial to inform a simulation model. The Screening for Haemoglobinopathies in First Trimester (SHIFT) trial.

By Dormandy E, Bryan S, Gulliford MC, Roberts T, Ades T, Calnan M, *et al.*

No. 21

Early referral strategies for management of people with markers of renal disease: a systematic review of the evidence of clinical effectiveness, cost-effectiveness and economic analysis.

By Black C, Sharma P, Scotland G, McCullough K, McGurn D, Robertson L, *et al.*

No. 22

A randomised controlled trial of cognitive behaviour therapy and motivational interviewing for people with Type 1 diabetes mellitus with persistent sub-optimal glycaemic control: A Diabetes and Psychological Therapies (ADaPT) study.

By Ismail K, Maissi E, Thomas S, Chalder T, Schmidt U, Bartlett J, *et al.*

No. 23

A randomised controlled equivalence trial to determine the effectiveness and cost-utility of manual chest physiotherapy techniques in the management of exacerbations of chronic obstructive pulmonary disease (MATREX).

By Cross J, Elender F, Barton G, Clark A, Shepstone L, Blyth A, *et al.*

No. 24

A systematic review and economic evaluation of the clinical effectiveness and cost-effectiveness of aldosterone antagonists for postmyocardial infarction heart failure.

By McKenna C, Burch J, Suekarran S, Walker S, Bakhai A, Witte K, *et al.*

No. 25

Avoiding and identifying errors in health technology assessment models: qualitative study and methodological review.

By Chilcott JB, Tappenden P, Rawdin A, Johnson M, Kaltenthaler E, Paisley S, *et al.*

No. 26

BoTULS: a multicentre randomised controlled trial to evaluate the clinical effectiveness and cost-effectiveness of treating upper limb spasticity due to stroke with botulinum toxin type A.

By Shaw L, Rodgers H, Price C, van Wijck F, Shackley P, Steen N, *et al.*, on behalf of the BoTULS investigators.



Health Technology Assessment programme

Director,
Professor Tom Walley,
Director, NIHR HTA
programme, Professor of
Clinical Pharmacology,
University of Liverpool

Deputy Director,
Professor Jon Nicholl,
Director, Medical Care Research
Unit, University of Sheffield

Prioritisation Strategy Group

Members

Chair,
Professor Tom Walley,
Director, NIHR HTA
programme, Professor of
Clinical Pharmacology,
University of Liverpool

Deputy Chair,
Professor Jon Nicholl,
Director, Medical Care Research
Unit, University of Sheffield

Dr Bob Coates,
Consultant Advisor, NETSCC,
HTA

Dr Andrew Cook,
Consultant Advisor, NETSCC,
HTA

Dr Peter Davidson,
Director of NETSCC, Health
Technology Assessment

Professor Robin E Ferner,
Consultant Physician and
Director, West Midlands Centre
for Adverse Drug Reactions,
City Hospital NHS Trust,
Birmingham

Professor Paul Glasziou,
Professor of Evidence-Based
Medicine, University of Oxford

Dr Nick Hicks,
Consultant Adviser, NETSCC,
HTA

Dr Edmund Jessop,
Medical Adviser, National
Specialist, National
Commissioning Group (NCG),
Department of Health, London

Ms Lynn Kerridge,
Chief Executive Officer,
NETSCC and NETSCC, HTA

Dr Ruairidh Milne,
Director of NETSCC External
Relations

Ms Kay Pattison,
Senior NIHR Programme
Manager, Department of
Health

Ms Pamela Young,
Specialist Programme Manager,
NETSCC, HTA

HTA Commissioning Board

Members

Programme Director,
Professor Tom Walley,
Director, NIHR HTA
programme, Professor of
Clinical Pharmacology,
University of Liverpool

Chairs,
Professor Sallie Lamb,
Director, Warwick Clinical Trials
Unit

Professor Hywel Williams,
Director, Nottingham Clinical
Trials Unit

Deputy Chair,
Dr Andrew Farmer,
Senior Lecturer in General
Practice, Department of
Primary Health Care,
University of Oxford

Professor Ann Ashburn,
Professor of Rehabilitation
and Head of Research,
Southampton General Hospital

Professor Deborah Ashby,
Professor of Medical Statistics,
Queen Mary, University of
London

Professor John Cairns,
Professor of Health Economics,
London School of Hygiene and
Tropical Medicine

Professor Peter Croft,
Director of Primary Care
Sciences Research Centre, Keele
University

Professor Nicky Cullum,
Director of Centre for Evidence-
Based Nursing, University of
York

Professor Jenny Donovan,
Professor of Social Medicine,
University of Bristol

Professor Steve Halligan,
Professor of Gastrointestinal
Radiology, University College
Hospital, London

Professor Freddie Hamdy,
Professor of Urology,
University of Sheffield

Professor Allan House,
Professor of Liaison Psychiatry,
University of Leeds

Dr Martin J Landray,
Reader in Epidemiology,
Honorary Consultant Physician,
Clinical Trial Service Unit,
University of Oxford

Professor Stuart Logan,
Director of Health & Social
Care Research, The Peninsula
Medical School, Universities of
Exeter and Plymouth

Dr Rafael Perera,
Lecturer in Medical Statistics,
Department of Primary Health
Care, University of Oxford

Professor Ian Roberts,
Professor of Epidemiology &
Public Health, London School
of Hygiene and Tropical
Medicine

Professor Mark Sculpher,
Professor of Health Economics,
University of York

Professor Helen Smith,
Professor of Primary Care,
University of Brighton

Professor Kate Thomas,
Professor of Complementary &
Alternative Medicine Research,
University of Leeds

Professor David John
Torgerson,
Director of York Trials Unit,
University of York

Observers

Ms Kay Pattison,
Section Head, NHS R&D
Programme, Department of
Health

Dr Morven Roberts,
Clinical Trials Manager,
Medical Research Council

Diagnostic Technologies and Screening Panel

Members

Chair,
Professor Paul Glasziou,
Professor of Evidence-Based
Medicine, University of Oxford

Deputy Chair,
Dr David Elliman,
Consultant Paediatrician and
Honorary Senior Lecturer,
Great Ormond Street Hospital,
London

Professor Judith E Adams,
Consultant Radiologist,
Manchester Royal Infirmary,
Central Manchester &
Manchester Children's
University Hospitals NHS
Trust, and Professor of
Diagnostic Radiology, Imaging
Science and Biomedical
Engineering, Cancer &
Imaging Sciences, University of
Manchester

Mr A S Arunkalaivanan,
Honorary Senior Lecturer,
University of Birmingham and
Consultant Urogynaecologist
and Obstetrician, City Hospital

Dr Dianne Baralle,
Consultant & Senior Lecturer
in Clinical Genetics, Human
Genetics Division & Wessex
Clinical Genetics Service,
Southampton, University of
Southampton

Dr Stephanie Dancer,
Consultant Microbiologist,
Hairmyres Hospital, East
Kilbride

Dr Ron Gray,
Consultant, National Perinatal
Epidemiology Unit, Institute of
Health Sciences, University of
Oxford

Professor Paul D Griffiths,
Professor of Radiology,
Academic Unit of Radiology,
University of Sheffield

Mr Martin Hooper,
Service User Representative

Professor Anthony Robert
Kendrick,
Professor of Primary
Medical Care, University of
Southampton

Dr Susanne M Ludgate,
Director, Medical Devices
Agency, London

Dr Anne Mackie,
Director of Programmes, UK
National Screening Committee

Dr David Mathew
Service User Representative

Dr Michael Millar, Lead
Consultant in Microbiology,
Department of Pathology &
Microbiology, Barts and The
London NHS Trust, Royal
London Hospital

Mr Stephen Pilling,
Director, Centre for Outcomes,
Research & Effectiveness,
University College London

Mrs Una Rennard,
Service User Representative

Ms Jane Smith,
Consultant Ultrasound
Practitioner, Ultrasound
Department, Leeds Teaching
Hospital NHS Trust, Leeds

Dr W Stuart A Smellie,
Consultant, Bishop Auckland
General Hospital

Professor Lindsay Wilson
Turnbull,
Scientific Director of the
Centre for Magnetic Resonance
Investigations and YCR
Professor of Radiology, Hull
Royal Infirmary

Dr Alan J Williams,
Consultant in General
Medicine, Department of
Thoracic Medicine, The Royal
Bournemouth Hospital

Observers

Dr Tim Elliott,
Team Leader, Cancer
Screening, Department of
Health

Dr Catherine Moody,
Programme Manager,
Neuroscience and Mental
Health Board

Dr Ursula Wells,
Principal Research Officer,
Department of Health

Disease Prevention Panel

Members

Chair,
Dr Edmund Jessop,
Medical Adviser, National
Specialist Commissioning
Advisory Group (NSCAG),
Department of Health

Deputy Chair,
**Professor Margaret
Thorogood,**
Professor of Epidemiology,
University of Warwick Medical
School, Coventry

Dr Robert Cook
Clinical Programmes Director,
Bazian Ltd, London

Dr Elizabeth Fellow-Smith,
Medical Director, West London
Mental Health Trust, Middlesex

Dr Colin Greaves
Senior Research Fellow,
Peninsular Medical School
(Primary Care)

Dr John Jackson,
General Practitioner, Parkway
Medical Centre, Newcastle
upon Tyne

Dr Russell Jago,
Senior Lecturer in Exercise,
Nutrition and Health, Centre
for Sport, Exercise and Health,
University of Bristol

Dr Chris McCall,
General Practitioner, The
Hadleigh Practice, Corfe
Mullen, Dorset

Miss Nicky Mullany,
Service User Representative

Dr Julie Mytton,
Locum Consultant in Public
Health Medicine, Bristol
Primary Care Trust

Professor Irwin Nazareth,
Professor of Primary Care
and Director, Department of
Primary Care and Population
Sciences, University College
London

Professor Ian Roberts,
Professor of Epidemiology and
Public Health, London School
of Hygiene & Tropical Medicine

Professor Carol Tannahill,
Glasgow Centre for Population
Health

Mrs Jean Thurston,
Service User Representative

Professor David Weller,
Head, School of Clinical
Science and Community
Health, University of
Edinburgh

Observers

Ms Christine McGuire,
Research & Development,
Department of Health

Ms Kay Pattison
Senior NIHR Programme
Manager, Department of
Health

Dr Caroline Stone,
Programme Manager, Medical
Research Council

External Devices and Physical Therapies Panel

Members

Chair, Dr John Pounsford, Consultant Physician North Bristol NHS Trust, Bristol	Dr Dawn Carnes, Senior Research Fellow, Barts and the London School of Medicine and Dentistry, London	Dr Peter Martin, Consultant Neurologist, Addenbrooke's Hospital, Cambridge	Dr Pippa Tyrrell, Stroke Medicine, Senior Lecturer/Consultant Stroke Physician, Salford Royal Foundation Hospitals' Trust, Salford
Deputy Chair, Professor E Andrea Nelson, Reader in Wound Healing and Director of Research, University of Leeds, Leeds	Dr Emma Clark, Clinician Scientist Fellow & Cons. Rheumatologist, University of Bristol, Bristol	Dr Lorraine Pinnigton, Associate Professor in Rehabilitation, University of Nottingham, Nottingham	Dr Sarah Tyson, Senior Research Fellow & Associate Head of School, University of Salford, Salford
Professor Bipin Bhakta Charterhouse Professor in Rehabilitation Medicine, University of Leeds, Leeds	Mrs Anthea De Barton-Watson, Service User Representative	Dr Kate Radford, Division of Rehabilitation and Ageing, School of Community Health Sciences, University of Nottingham, Nottingham	Dr Nefyn Williams, Clinical Senior Lecturer, Cardiff University, Cardiff
Mrs Penny Calder Service User Representative	Professor Christopher Griffiths, Professor of Primary Care, Barts and the London School of Medicine and Dentistry, London	Mr Jim Reece, Service User Representative	
Professor Paul Carding, Professor of Voice Pathology, Newcastle Hospital NHS Trust, Newcastle	Dr Shaheen Hamdy, Clinical Senior Lecturer and Consultant Physician, University of Manchester, Manchester	Professor Maria Stokes, Professor of Neuromusculoskeletal Rehabilitation, University of Southampton, Southampton	

Observers

Dr Phillip Leech, Principal Medical Officer for Primary Care, Department of Health, London	Ms Kay Pattison Senior NIHR Programme Manager, Department of Health	Dr Morven Roberts, Clinical Trials Manager, MRC, London	Dr Ursula Wells PRP, DH, London
---	--	---	------------------------------------

Interventional Procedures Panel

Members

Chair, Professor Jonathan Michaels, Consultant Surgeon & Honorary Clinical Lecturer, University of Sheffield	Mr Seamus Eckford, Consultant in Obstetrics & Gynaecology, North Devon District Hospital	Dr Nadim Malik, Consultant Cardiologist/ Honorary Lecturer, University of Manchester	Dr Ashish Paul, Medical Director, Bedfordshire PCT
Mr David P Britt, Service User Representative, Cheshire	Professor David Taggart, Consultant Cardiothoracic Surgeon, John Radcliffe Hospital	Mr Hisham Mehanna, Consultant & Honorary Associate Professor, University Hospitals Coventry & Warwickshire NHS Trust	Dr Sarah Purdy, Consultant Senior Lecturer, University of Bristol
Mr Sankaran ChandraSekharan, Consultant Surgeon, Colchester Hospital University NHS Foundation Trust	Dr Matthew Hatton, Consultant in Clinical Oncology, Sheffield Teaching Hospital Foundation Trust	Dr Jane Montgomery, Consultant in Anaesthetics and Critical Care, South Devon Healthcare NHS Foundation Trust	Mr Michael Thomas, Consultant Colorectal Surgeon, Bristol Royal Infirmary
Professor Nicholas Clarke, Consultant Orthopaedic Surgeon, Southampton University Hospitals NHS Trust	Dr John Holden, General Practitioner, Garswood Surgery, Wigan	Dr Simon Padley, Consultant Radiologist, Chelsea & Westminster Hospital	Professor Yit Chiun Yang, Consultant Ophthalmologist, Royal Wolverhampton Hospitals NHS Trust
			Mrs Isabel Boyer, Service User Representative, London

Pharmaceuticals Panel

Members

Chair,

Professor Imti Choonara,
Professor in Child Health,
University of Nottingham

Deputy Chair,

Dr Lesley Wise,
Unit Manager,
Pharmacoepidemiology
Research Unit, VRMM,
Medicines & Healthcare
Products Regulatory Agency

Mrs Nicola Carey,
Senior Research Fellow,
School of Health and Social
Care, The University of
Reading

Mr John Chapman,
Service User Representative

Dr Peter Elton,
Director of Public Health,
Bury Primary Care Trust

Professor Robin Ferner,
Consultant Physician and
Director, West Midlands Centre
for Adverse Drug Reactions,
City Hospital NHS Trust,
Birmingham

Dr Ben Goldacre,
Research Fellow, Division of
Psychological Medicine and
Psychiatry, King's College
London

Dr Bill Gutteridge,
Medical Adviser, London
Strategic Health Authority

Dr Dyfrig Hughes,
Reader in Pharmacoeconomics
and Deputy Director, Centre
for Economics and Policy in
Health, IMSCaR, Bangor
University

Dr Yoon K Loke,
Senior Lecturer in Clinical
Pharmacology, University of
East Anglia

Professor Femi Oyeboode,
Consultant Psychiatrist
and Head of Department,
University of Birmingham

Dr Andrew Prentice,
Senior Lecturer and Consultant
Obstetrician and Gynaecologist,
The Rosie Hospital, University
of Cambridge

Dr Martin Shelly,
General Practitioner, Leeds,
and Associate Director, NHS
Clinical Governance Support
Team, Leicester

Dr Gillian Shepherd,
Director, Health and Clinical
Excellence, Merck Serono Ltd

Mrs Katrina Simister,
Assistant Director New
Medicines, National Prescribing
Centre, Liverpool

Mr David Symes,
Service User Representative

Observers

Ms Kay Pattison
Senior NIHR Programme
Manager, Department of
Health

Mr Simon Reeve,
Head of Clinical and Cost-
Effectiveness, Medicines,
Pharmacy and Industry Group,
Department of Health

Dr Heike Weber,
Programme Manager,
Medical Research Council

Dr Ursula Wells,
Principal Research Officer,
Department of Health

Psychological and Community Therapies Panel

Members

Chair,

Professor Scott Weich,
Professor of Psychiatry,
University of Warwick

Professor Jane Barlow,
Professor of Public Health in
the Early Years, Health Sciences
Research Institute, Warwick
Medical School

Dr Sabyasachi Bhaumik,
Consultant Psychiatrist,
Leicestershire Partnership NHS
Trust

Mrs Val Carlill,
Service User Representative,
Gloucestershire

Dr Steve Cunningham,
Consultant Respiratory
Paediatrician, Lothian Health
Board

Dr Anne Hesketh,
Senior Clinical Lecturer in
Speech and Language Therapy,
University of Manchester

Dr Yann Lefevre,
GP Partner, Burrage Road
Surgery, London

Dr Jeremy J Murphy,
Consultant Physician &
Cardiologist, County Durham &
Darlington Foundation Trust

Mr John Needham,
Service User, Buckinghamshire

Ms Mary Nettle,
Mental Health User Consultant,
Gloucestershire

Professor John Potter,
Professor of Ageing and Stroke
Medicine, University of East
Anglia

Dr Greta Rait,
Senior Clinical Lecturer and
General Practitioner, University
College London

Dr Paul Ramchandani,
Senior Research Fellow/Cons.
Child Psychiatrist, University of
Oxford

Dr Howard Ring,
Consultant & University
Lecturer in Psychiatry,
University of Cambridge

Dr Karen Roberts,
Nurse/Consultant, Dunston Hill
Hospital, Tyne and Wear

Dr Karim Saad,
Consultant in Old Age
Psychiatry, Coventry &
Warwickshire Partnership Trust

Dr Alastair Sutcliffe,
Senior Lecturer, University
College London

Dr Simon Wright,
GP Partner, Walkden Medical
Centre, Manchester

Observers

Ms Kay Pattison
Senior NIHR Programme
Manager, Department of
Health

Dr Morven Roberts,
Clinical Trials Manager, MRC,
London

Professor Tom Walley,
HTA Programme Director,
Liverpool

Dr Ursula Wells,
Policy Research Programme,
DH, London

Expert Advisory Network

Members

Professor Douglas Altman,
Professor of Statistics in
Medicine, Centre for Statistics
in Medicine, University of
Oxford

Professor John Bond,
Professor of Social Gerontology
& Health Services Research,
University of Newcastle upon
Tyne

Professor Andrew Bradbury,
Professor of Vascular Surgery,
Solihull Hospital, Birmingham

Mr Shaun Brogan,
Chief Executive, Ridgeway
Primary Care Group, Aylesbury

Mrs Stella Burnside OBE,
Chief Executive, Regulation
and Improvement Authority,
Belfast

Ms Tracy Bury,
Project Manager, World
Confederation for Physical
Therapy, London

Professor Iain T Cameron,
Professor of Obstetrics and
Gynaecology and Head of the
School of Medicine, University
of Southampton

Dr Christine Clark,
Medical Writer and Consultant
Pharmacist, Rossendale

Professor Collette Clifford,
Professor of Nursing and
Head of Research, The
Medical School, University of
Birmingham

Professor Barry Cookson,
Director, Laboratory of Hospital
Infection, Public Health
Laboratory Service, London

Dr Carl Counsel, Clinical Senior Lecturer in
Neurology, University of
Aberdeen

Professor Howard Cuckle,
Professor of Reproductive
Epidemiology, Department
of Paediatrics, Obstetrics &
Gynaecology, University of
Leeds

Dr Katherine Darton,
Information Unit, MIND – The
Mental Health Charity, London

Professor Carol Dezateux,
Professor of Paediatric
Epidemiology, Institute of Child
Health, London

Mr John Dunning,
Consultant Cardiothoracic
Surgeon, Papworth Hospital
NHS Trust, Cambridge

Mr Jonothan Earnshaw,
Consultant Vascular Surgeon,
Gloucestershire Royal Hospital,
Gloucester

Professor Martin Eccles,
Professor of Clinical
Effectiveness, Centre for Health
Services Research, University of
Newcastle upon Tyne

Professor Pam Enderby,
Dean of Faculty of Medicine,
Institute of General Practice
and Primary Care, University of
Sheffield

Professor Gene Feder,
Professor of Primary Care
Research & Development,
Centre for Health Sciences,
Barts and The London School
of Medicine and Dentistry

Mr Leonard R Fenwick,
Chief Executive, Freeman
Hospital, Newcastle upon Tyne

Mrs Gillian Fletcher,
Antenatal Teacher and Tutor
and President, National
Childbirth Trust, Henfield

Professor Jayne Franklyn,
Professor of Medicine,
University of Birmingham

Mr Tam Fry,
Honorary Chairman, Child
Growth Foundation, London

Professor Fiona Gilbert,
Consultant Radiologist and
ACRN Member, University of
Aberdeen

Professor Paul Gregg,
Professor of Orthopaedic
Surgical Science, South Tees
Hospital NHS Trust

Bec Hanley,
Co-director, TwoCan Associates,
West Sussex

Dr Maryann L Hardy,
Senior Lecturer, University of
Bradford

Mrs Sharon Hart,
Healthcare Management
Consultant, Reading

Professor Robert E Hawkins,
CRC Professor and Director
of Medical Oncology, Christie
CRC Research Centre,
Christie Hospital NHS Trust,
Manchester

Professor Richard Hobbs,
Head of Department of Primary
Care & General Practice,
University of Birmingham

Professor Alan Horwich,
Dean and Section Chairman,
The Institute of Cancer
Research, London

Professor Allen Hutchinson,
Director of Public Health and
Deputy Dean of ScHARR,
University of Sheffield

Professor Peter Jones,
Professor of Psychiatry,
University of Cambridge,
Cambridge

Professor Stan Kaye,
Cancer Research UK Professor
of Medical Oncology, Royal
Marsden Hospital and Institute
of Cancer Research, Surrey

Dr Duncan Keeley,
General Practitioner (Dr Burch
& Ptms), The Health Centre,
Thame

Dr Donna Lamping,
Research Degrees Programme
Director and Reader in
Psychology, Health Services
Research Unit, London School
of Hygiene and Tropical
Medicine, London

Mr George Levy,
Chief Executive, Motor
Neurone Disease Association,
Northampton

Professor James Lindesay,
Professor of Psychiatry for the
Elderly, University of Leicester

Professor Julian Little,
Professor of Human Genome
Epidemiology, University of
Ottawa

Professor Alistaire McGuire,
Professor of Health Economics,
London School of Economics

Professor Rajan Madhok,
Medical Director and Director
of Public Health, Directorate
of Clinical Strategy & Public
Health, North & East Yorkshire
& Northern Lincolnshire
Health Authority, York

Professor Alexander Markham,
Director, Molecular Medicine
Unit, St James's University
Hospital, Leeds

Dr Peter Moore,
Freelance Science Writer,
Ashted

Dr Andrew Mortimore,
Public Health Director,
Southampton City Primary
Care Trust

Dr Sue Moss,
Associate Director, Cancer
Screening Evaluation Unit,
Institute of Cancer Research,
Sutton

Professor Miranda Mugford,
Professor of Health Economics
and Group Co-ordinator,
University of East Anglia

Professor Jim Neilson,
Head of School of Reproductive
& Developmental Medicine
and Professor of Obstetrics
and Gynaecology, University of
Liverpool

Mrs Julietta Patnick,
National Co-ordinator, NHS
Cancer Screening Programmes,
Sheffield

Professor Robert Peveler,
Professor of Liaison Psychiatry,
Royal South Hants Hospital,
Southampton

Professor Chris Price,
Director of Clinical Research,
Bayer Diagnostics Europe,
Stoke Poges

Professor William Rosenberg,
Professor of Hepatology
and Consultant Physician,
University of Southampton

Professor Peter Sandercock,
Professor of Medical Neurology,
Department of Clinical
Neurosciences, University of
Edinburgh

Dr Susan Schonfield,
Consultant in Public Health,
Hillingdon Primary Care Trust,
Middlesex

Dr Eamonn Sheridan,
Consultant in Clinical Genetics,
St James's University Hospital,
Leeds

Dr Margaret Somerville,
Director of Public Health
Learning, Peninsula Medical
School, University of Plymouth

Professor Sarah Stewart-Brown,
Professor of Public Health,
Division of Health in the
Community, University of
Warwick, Coventry

Professor Ala Szczepura,
Professor of Health Service
Research, Centre for Health
Services Studies, University of
Warwick, Coventry

Mrs Joan Webster,
Consumer Member, Southern
Derbyshire Community Health
Council

Professor Martin Whittle,
Clinical Co-director, National
Co-ordinating Centre for
Women's and Children's
Health, Lymington

Feedback

The HTA programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (www.hta.ac.uk) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.