

# **An assessment of the impact of the NHS Health Technology Assessment Programme**

S Hanney, M Buxton, C Green, D Coulson  
and J Raftery



December 2007

**Health Technology Assessment  
NHS R&D HTA Programme**  
[www.hta.ac.uk](http://www.hta.ac.uk)





**INAHTA**

### **How to obtain copies of this and other HTA Programme reports.**

An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (<http://www.hta.ac.uk>). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs 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 monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

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

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

### **Contact details are as follows:**

HTA Despatch  
c/o Direct Mail Works Ltd  
4 Oakwood Business Centre  
Downley, HAVANT PO9 2NP, UK

Email: [orders@hta.ac.uk](mailto:orders@hta.ac.uk)  
Tel: 02392 492 000  
Fax: 02392 478 555  
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

### **Payment methods**

#### *Paying by cheque*

If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

#### *Paying by credit card*

The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

#### *Paying by official purchase order*

You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

### **How do I get a copy of HTA on CD?**

Please use the form on the HTA website ([www.hta.ac.uk/htacd.htm](http://www.hta.ac.uk/htacd.htm)). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

---

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

# An assessment of the impact of the NHS Health Technology Assessment Programme

S Hanney,<sup>1</sup> M Buxton,<sup>1</sup> C Green,<sup>2</sup> D Coulson<sup>2</sup>  
and J Raftery<sup>3\*†</sup>

<sup>1</sup> Health Economics Research Group, Brunel University, Uxbridge, UK

<sup>2</sup> NCCHTA, University of Southampton, UK

<sup>3</sup> Department of Epidemiology and Public Health, University of Birmingham, UK

\* Corresponding author

† Present address: Wessex Institute of Health Research and Development, University of Southampton, UK

**Declared competing interests of authors:** see page iv

Published December 2007

---

This report should be referenced as follows:

Hanney S, Buxton M, Green C, Coulson D, Raftery J. An assessment of the impact of the NHS Health Technology Assessment Programme. *Health Technol Assess* 2007; **11**(53).

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

# NIHR Health Technology Assessment Programme

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

Secondly, 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.

Thirdly, 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 monograph series *Health Technology Assessment*.

## Criteria for inclusion in the HTA monograph series

Reports are published in the HTA monograph 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 monograph was commissioned by the HTA Programme as project number 03/67/01. The contractual start date was in April 2005. The draft report began editorial review in October 2006 and was accepted for publication in May 2007. As the funder, by devising a commissioning brief, the HTA Programme 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

Series Editors:

Dr Aileen Clarke, Dr Peter Davidson, Dr Chris Hyde,  
Dr John Powell, Dr Rob Riemsma and Professor Ken Stein

Programme Managers:

Sarah Llewellyn Lloyd, Stephen Lemon, Kate Rodger,  
Stephanie Russell and Pauline Swinburne

ISSN 1366-5278

© Queen's Printer and Controller of HMSO 2007

This monograph 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: NCCHTA, Mailpoint 728, Boldrewood, University of Southampton, Southampton, SO16 7PX, UK.

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.

Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.



## Abstract

### An assessment of the impact of the NHS Health Technology Assessment Programme

S Hanney,<sup>1</sup> M Buxton,<sup>1</sup> C Green,<sup>2</sup> D Coulson<sup>2</sup> and J Raftery<sup>3\*†</sup>

<sup>1</sup> Health Economics Research Group, Brunel University, Uxbridge, UK

<sup>2</sup> NCCHTA, University of Southampton, UK

<sup>3</sup> Department of Epidemiology and Public Health, University of Birmingham, UK

\* Corresponding author

† Present address: Wessex Institute of Health Research and Development, University of Southampton, UK

**Objectives:** To consider how the impact of the NHS Health Technology Assessment (HTA) Programme should be measured. To determine what models are available and their strengths and weaknesses. To assess the impact of the first 10 years of the NHS HTA programme from its inception in 1993 to June 2003 and to identify the factors associated with HTA research that are making an impact.

**Data sources:** Main electronic databases from 1990 to June 2005. The documentation of the National Coordinating Centre for Health Technology Assessment (NCCHTA). Questionnaires to eligible researchers. Interviews with lead investigators. Case study documentation.

**Review methods:** A literature review of research programmes was carried out, the work of the NCCHTA was reviewed, lead researchers were surveyed and 16 detailed case studies were undertaken. Each case study was written up using the payback framework. A cross-case analysis informed the analysis of factors associated with achieving payback. Each case study was scored for impact before and after the interview to assess the gain in information due to the interview. The draft write-up of each study was checked with each respondent for accuracy and changed if necessary.

**Results:** The literature review identified a highly diverse literature but confirmed that the 'payback' framework pioneered by Buxton and Hanney was the most widely used and most appropriate model available. The review also confirmed that impact on knowledge generation was more easily quantified than that on policy, behaviour or especially health gain. The review of the included studies indicated a higher level of impact on policy than is often assumed to occur. The survey showed that data pertinent to payback exist and can be collected. The completed questionnaires showed that the HTA Programme had considerable impact in terms of publications, dissemination, policy and behaviour. It also

showed, as expected, that different parts of the Programme had different impacts. The Technology Assessment Reports (TARs) for the National Institute for Health and Clinical Excellence (NICE) had the clearest impact on policy in the form of NICE guidance. Mean publications per project were 2.93 (1.98 excluding the monographs), above the level reported for other programmes. The case studies revealed the large diversity in the levels and forms of impacts and the ways in which they arise. All the NICE TARs and more than half of the other case studies had some impact on policy making at the national level whether through NICE, the National Screening Committee, the National Service Frameworks, professional bodies or the Department of Health. This underlines the importance of having a customer or 'receptor' body. A few case studies had very considerable impact in terms of knowledge production and in informing national and international policies. In some of these the principal investigator had prior expertise and/or a research record in the topic. The case studies confirmed the questionnaire responses but also showed how some projects led to further research.

**Conclusions:** This study concluded that the HTA Programme has had considerable impact in terms of knowledge generation and perceived impact on policy and to some extent on practice. This high impact may have resulted partly from the HTA Programme's objectives, in that topics tend to be of relevance to the NHS and have policy customers. The required use of scientific methods, notably systematic reviews and trials, coupled with strict peer reviewing, may have helped projects publish in high-quality peer-reviewed journals. Further research should cover more detailed, comprehensive case studies, as well as enhancement of the 'payback framework'. A project that collated health research impact studies in an ongoing manner and analysed them in a consistent fashion would also be valuable.

## Declared competing interests of authors

None of the authors had any direct conflict of interest, financial or otherwise. As indirect conflicts of interest could be inferred due to the positions several of the authors occupy, these are outlined below.

One source of indirect conflict of interest could be due to employment of some members of the team by NCCHTA. NCCHTA commissioned this project from the Wessex Institute for Health R&D (WIHRD) under its direct commissioning facility. WIHRD is the academic department within which NCCHTA is based.

James Raftery became Director of NCCHTA on a part-time basis from May 2005. As Director he might be inferred to have an interest in the evaluation of the HTA Programme being shown in best light. JR is also Director of WIHRD. As part of his academic duties in WIHRD, JR took responsibility for managing the overall project, analysing the survey results and drafting various chapters as indicated below. The original lead researcher, John Powell, had moved in early 2005 to a new post in Warwick University. As the project had been commissioned in 2004, JR had no role in the original design of the project. His input was largely confined to analysis of the survey results and drafting Chapters 4, 5 and 7.

Christine Solomon, who organised the survey, was employed part time by NCCHTA, part time by WIHRD working on this project from 2004 to 2005. She had no role in the analysis of results or

writing this report. Colin Green, employed by WIHRD, also worked part time for NCCHTA. His role was confined to the literature review and commenting on the draft report. Neither CS nor CG, both of whom have since left WIHRD employment, could be inferred to gain from a positive evaluation of the HTA Programme.

A second source of indirect conflict of interest could be thought to arise for Steve Hanney and Martin Buxton who, as originators and exponents of the 'payback framework', would wish to see the use of that framework in this project, as in any other applications.

As indicated in the report, the decision to use the payback framework reflected a recommendation in the project brief based on an exploratory study led by Ruairidh Milne. This decision was endorsed by the external Advisory Group before the literature review was complete. In fact, the literature review showed that relatively few alternative frameworks existed and that 'payback' had been most often applied and was the most appropriate. The involvement of CG with SH in the literature review ensured that decisions on which studies to include were unbiased.

As academics and professional researchers, the authors do not believe they have allowed any bias, whether due to the factors described above or otherwise, to affect the design of the work, its analysis or the conclusions drawn. However, they also recognise that the test of this is up to the reader.



# Contents

<b>List of abbreviations</b> .....	vii	Ethics approval .....	48
<b>Executive summary</b> .....	ix	<b>4 The NHS HTA Programme</b> .....	49
<b>I Introduction</b> .....	1	Introduction .....	49
Background .....	1	Structure .....	49
The importance of assessing the impact of research and development and health technology assessment .....	1	Types of research .....	49
Main research questions .....	2	Number of projects .....	50
Exploratory study .....	2	Monitoring .....	50
Project team .....	3	Assessing the impact of the HTA Programme .....	51
The NHS HTA Programme .....	3	<b>5 Survey of lead researchers</b> .....	53
Structure of the report .....	4	Introduction .....	53
<b>2 An assessment of the impact of the NHS HTA Programme: literature review</b> .....	5	Exclusions .....	53
Introduction .....	5	Questionnaire .....	53
Research questions .....	5	Results .....	53
Objectives .....	5	Conclusions .....	58
Research methods .....	5	<b>6 Impact of health technology assessments: 16 case studies</b> .....	59
Databases searched .....	5	Introduction .....	59
Inclusion/exclusion criteria (as per project protocol) .....	6	Methods .....	59
Literature identified .....	6	Selection of case studies .....	61
Review methods .....	7	Results .....	61
Quality criteria .....	7	Discussion and cross-case analysis .....	62
Review findings .....	7	Summary and conclusions .....	67
Summary of the literature identified .....	7	<b>7 Summary and discussion</b> .....	69
Techniques used in empirical assessment of the impact of health research programmes .....	8	Summary of findings .....	69
Conceptual or methodological approaches identified for the assessment of the impact of health research programmes .....	28	Discussion .....	72
An introduction to the empirical literature on assessment of health research programmes .....	34	<b>Acknowledgements</b> .....	77
Summary of key findings from studies that assess the impact of entire multi-project programmes .....	34	<b>References</b> .....	79
Discussion .....	36	<b>Appendix 1</b> Key references identified from literature search .....	83
Summary and conclusions .....	43	<b>Appendix 2</b> Literature review ‘evidence base’ .....	89
<b>3 Evaluative framework for this study</b> .....	45	<b>Appendix 3</b> Template for HTA review of literature on assessing the impact of research programmes .....	91
Previous use of the payback framework .....	45	<b>Appendix 4</b> Review of empirical literature/studies (by approach) .....	93
Aspects excluded from our assessment .....	48	<b>Appendix 5</b> Total hits for HTA projects in survey .....	109

<b>Appendix 6</b> HTA reports 1997–2006 considered by the National Screening Committee (NSC) .....	111
<b>Appendix 7</b> Questionnaire used in survey .....	115
<b>Appendix 8</b> Sixteen case studies .....	121
<b>Appendix 9</b> Scoring and re-scoring of the case study projects .....	171
<b>Health Technology Assessment reports published to date</b> .....	181
<b>Health Technology Assessment Programme</b> .....	197





## List of abbreviations

AHFMR	Alberta Heritage Foundation for Medical Research	HSR	health services research
AHRQ	Agency for Healthcare and Research Quality	ICER	incremental cost-effectiveness ratio
ALS	amyotrophic lateral sclerosis	IFN- $\alpha$	interferon-alpha
ARC	Arthritis Research Campaign	ISI	Institute for Scientific Information
BACP	British Association for Counselling and Psychotherapy	MAb	monoclonal antibody
BMT	bone marrow transplantation	MND	motor neurone disease
BPH	benign prostatic hyperplasia	MRC	Medical Research Council
CAD	computer-aided detection	MTA	Medical Technologies Administration
CBA	cost-benefit analysis	NCCHTA	National Coordinating Centre for Health Technology Assessment
CETS	Quebec Council on Health Care Technology Assessment	NCCSDO	National Coordinating Centre for Service Delivery and Organisation
CGS	Community Group Support	NCT	National Childbirth Trust
CI	confidence interval	NHSBSP	NHS Breast Screening Programme
CML	chronic myeloid leukaemia	NHTAP	National Health Technology Advisory Panel
CR	cytogenetic response	NICE	National Institute for Health and Clinical Excellence
DO	detrusor overactivity	NIH	National Institutes of Health
ESRC	Economic and Social Research Council	NIHR	National Institute for Health Research
FASEB	Federation of American Societies for Experimental Biology	NINDS	National Institute of Neurological Disorders and Stroke
FH	familial hypercholesterolaemia	NLHS	National List of Health Services
GPRA	Government Performance and Results Act	NSC	National Screening Committee
HERG	Health Economics Research Group	NUD	non-ulcer dyspepsia

*continued*

**List of abbreviations continued**

OER	Outcomes and Effectiveness Research	ScHARR	School of Health and Related Research
PCC	primary and community care	SHV	support health visitor
PenTAG	Peninsular Technology Assessment Group	SIGN	Scottish Intercollegiate Guidelines Network
PHRDC	Public Health Research and Development Committee	SMD	standardised mean difference
PROTO	Prostate Trials Office	SSFH	Social Support and Family Health
QALY	quality-adjusted life-year	TAR	Technology Assessment Report
QUADAS	Quality Assessment of Diagnostic Studies	TURP	transurethral resection of the prostate
RCP	Royal College of Physicians	TUVP	transurethral vaporisation of the prostate
RCT	randomised controlled trial	USI	urodynamic stress incontinence
RR	relative risk	WIHRD	Wessex Institute for Health Research and Development

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 at the end of the table.



## Executive summary

### Objectives

This project aimed to address two sets of questions:

1. How should the impact of the NHS Health Technology Assessment (HTA) Programme be measured? What models are available and what are their strengths and weaknesses?
2. What has been the impact of the first 10 years of the NHS HTA programme from its inception in 1993 to June 2003? What factors seem to be associated with HTA research making an impact?

### Methods

The first question was answered by a literature review of assessments of research programmes. Using standard techniques, an initial 1600 papers were identified. About 200 papers were put on a preliminary list and 46 studies were reviewed in detail. The review identified the methods used (desk, questionnaire, interviews and case studies), the main models or conceptual frameworks applied, and assessed strengths and weaknesses of the main approaches.

The answers to the second set of questions were sought using a 'multiple methods' approach, which triangulated National Coordinating Centre for Health Technology Assessment (NCCHTA) documentation, a survey of lead researchers and detailed case studies. The survey, using an established questionnaire, covered 204 eligible researchers who had led a project completed between 1993 and 2003.

Sixteen case studies provided more detailed examples of impact, the factors associated with impact and the best methods to assess impact. The 16 comprised nine clinical trials, four evidence synthesis research projects and three Technology Assessment Reports (TARs) for the National Institute for Health and Clinical Excellence (NICE). These were selected by stratified random sampling, to the best of our knowledge for the first time in payback analysis. The case studies consisted of interviews with lead investigators;

analysis of relevant documents including the main published papers and reports; and any other relevant reviews. Interviewees were asked to identify factors linked to the level of impact achieved. Each case study was written up using the payback framework. A cross-case analysis informed the analysis of factors associated with achieving payback.

Each case study was scored for impact before and after the interview to assess the gain in information due to the interview. The draft write-up of each study was checked with each respondent for accuracy and changed if necessary.

### Results

The literature review identified a highly diverse literature, but confirmed that the 'payback' framework pioneered by Buxton and Hanney was the most widely used and most appropriate model available. It encompassed key elements of many of the alternatives. The review confirmed that impact on knowledge generation was more easily quantified than that on policy, behaviour or especially health gain. The review of the included studies indicated a higher level of impact on policy than is often assumed to occur.

The diverse literature suggested that two different sets of studies might provide the most appropriate comparators for the two main parts of the NHS HTA Programme. Studies of the impact of 'HTA Programmes' for policy-making bodies can best be compared with the TARs for NICE. The group 'Other Health Research Programmes' provides comparisons for the primary and secondary research projects of the NHS HTA Programme.

The survey showed that data pertinent to payback exist and can be collected. However over one-third of projects did not respond, despite repeated reminders. Against this, a 100% response was not feasible as over the 10-year period several lead researchers had died, retired, moved or were otherwise not reachable.

The completed questionnaires confirmed, corrected and extended the data collated by

NCCHTA on publications and other indicators. They showed that the HTA Programme had considerable impact in terms of publications, dissemination, policy and behaviour. It also showed, as expected, that different parts of the Programme had different impacts. The TARs for NICE had the clearest impact on policy in the form of NICE guidance. Other policy 'customers' included the National Screening Committee (NSC) and National Service Frameworks.

Overall impacts measured in the survey were consistent with or somewhat better than those for other programmes identified in the literature review. Mean publications per project were 2.93 (1.98 excluding the monographs), above the level reported for other programmes. The proportion of NICE TARs reporting an impact on past policy at 96% was among the highest for the 'HTA Programmes'. The 60% of primary and secondary studies reporting an impact on policy was above the other programmes in its group (although some of the latter were responsive mode programmes which would not have been expected to make much impact on policy). The percentage of primary and secondary projects reporting an impact on behaviour was somewhere in the middle of the range for the 'Other Health Research Programmes'. Comparisons with other programmes must be treated with considerable caution due to differences in programme objectives, topics researched and methods of assessing impact.

The NCCHTA's reliance on researchers to inform it of publications was shown to lead to incomplete data. Around one-quarter of publications in peer-reviewed journals were missed. These data could probably be better collected using the Internet and then asking the researchers to correct and amend the resulting list. Other data such as those on presentations and further research could only be collected from the researchers.

The case studies revealed the large diversity in the levels and forms of impacts and the ways in which they arise. All the NICE TARs and more than half of the other case studies had some impact on policy making at the national level, whether through NICE, NSC, National Service Frameworks, professional bodies or the Department of Health. This underlines the importance of having a customer or 'receptor' body. A few case studies had very considerable impact in terms of knowledge production and in informing national and international policies. In some of these the Principal Investigator had prior

expertise and/or a research record in the topic. The case studies confirmed the questionnaire responses, but also provided more details, including information on how some projects led to further research.

All but one of the case studies with high impact had successful peer-reviewed publications and engaged in active dissemination. Although researchers were generally satisfied with NCCHTA, some complained about lengthy procedures and one about changes in study design.

The pre- and post-interview scoring showed reasonable correlations and high inter-rater reliability. This indicated that most researchers were not making exaggerated claims for impact in their questionnaire responses.

## Conclusions

This study concluded that the HTA Programme has had considerable impact in terms of knowledge generation and perceived impact on policy and to some extent on practice. This high impact may have resulted partly from the HTA Programme's objectives, in that topics tend to be of relevance to the NHS and have policy customers. The required use of scientific methods, notably systematic reviews and trials, coupled with strict peer reviewing, may have helped projects publish in high-quality peer-reviewed journals.

## Implications for healthcare

It could be argued on the basis of the review that the NHS would benefit from an expansion of the HTA Programme, and that more should be done to encourage NHS customers to seek research relevant to their own work with a view to changing practice.

## Implications for the HTA Programme

Recommendations were made on how the HTA Programme could improve, including:

- more rapid commissioning and publication
- improved collection of data on publications
- better archiving of documents relating to the origination, development and commissioning of research projects
- minimising late changes in research project design
- increased feedback to researchers on the dissemination of their reports, particularly on the large numbers of web views and downloads.

## Recommendations for research

Three main areas for further research were identified:

1. More detailed, comprehensive case studies, based on the payback approach, should be undertaken on selective projects within research programmes.
2. Further enhancement of the 'payback framework' would be useful. The impact

- assessment provided lessons on how the payback framework might be improved, specifically how the questionnaire might be amended in the light of the case studies, and how factors associated with high impact can be further refined.
3. A project that collated health research impact studies in an ongoing manner and analysed them in a consistent fashion would be valuable.



# Chapter I

## Introduction

### Background

Health technology assessment describes the systematic and rigorous application of scientific methods to the evaluation of healthcare technologies. The aim of health technology assessment is to provide evidence to guide healthcare decisions and policy making. Health technology assessment focuses on the effectiveness, appropriateness and cost of technologies. It asks four fundamental questions: (1) does the intervention work?; (2) for whom?; (3) at what cost?; (4) how does it compare with the alternatives?<sup>1</sup>

Health technology assessment encourages a critical approach that seeks evidence of the benefit provided by a technology and typically assesses on the balance of that benefit against the opportunity costs of the resources utilised. It is particularly apposite, therefore, that funders of health technology assessment should apply this critical, evaluative approach to health technology assessment itself and to assess its benefits and costs. Proponents of health technology assessment should no more expect that the benefits of health technology assessment be taken as self-evident, than should the proponents of any other health-related technology.

However, the assessment of the benefits or impact from health technology assessment-type research is not necessarily easy or uncontentious, although it is less problematic than assessing the impact from research of a much more basic nature.

This report presents the findings of an assessment of the impact of the NHS Health Technology Assessment (HTA) Programme commissioned by the Programme in 2004. It provides a critical review of methods for assessing such research and of existing studies of relevant programmes and also a quantitative and qualitative assessment of projects funded by the Programme. The empirical assessment further contributes to the understanding of methods by comparing assessments based on alternative methods/sources of information.

### The importance of assessing the impact of research and development and health technology assessment

The need to demonstrate the value of research is increasingly being recognised and acted upon. The recently published report from the UK Evaluation Forum – a joint initiative from the Academy of Medical Sciences, Medical Research Council and Wellcome Trust – is the most recent of many manifestations of this recognition.<sup>2</sup>

The most common reason for such evaluation is as a form of advocacy for research: to justify existing levels of, and desired increases in, spending on such programmes. The Evaluation Forum report noted the overt advocacy role of ‘Research! America’, which has an explicit mission to make medical and health research a much higher national priority in the USA. Other bodies may be less overt in their mission but they still have an understandable interest in such positive support for their work. Key studies attempting to estimate the rate of return on medical research in the USA were published eventually as an academic monograph,<sup>3</sup> but were commissioned and presented at a conference organised by the Lasker Foundation’s Funding First programme, which is committed to building and publicising the case for a sustained, long-term national commitment to medical research (URL: <http://www.laskerfoundation.org/about/ffirst.html>, accessed 8 June 2006).

The second need is related but driven by outside forces: the general requirements or specific institutional requirements for accountability. Thus work on impact in the USA by the National Institutes of Health, for example, has been driven by the requirements of the US Government Performance and Results Act of 1993. Current efforts by the Canadian Institutes of Health Research can be linked to the Report of the Auditor General of Canada (1994), which concluded that departments and agencies should establish mechanisms and practices to

demonstrate the results of their science and technology activities. In the UK, the National Audit Office's conclusion that departments "have no systematic mechanisms for measuring the overall impact of their research effort" post-dates the genesis of this report but does partly explain the current level of UK interest.<sup>4</sup>

The third argument for such work is that it should inform, and make more efficient the funding, commissioning and undertaking of research. The evidence base for funding policies is currently very weak, and views as to what works best and which types of projects have most impact are largely based on impressionistic knowledge rather than systematic enquiry. A key driver for the study on the returns from arthritis research commissioned by the Arthritis Research Campaign was to inform its future funding decisions by providing a better understanding of the relative impacts of different funding streams.<sup>5</sup> Indeed, one of the key issues addressed in the exploratory study by the National Coordinating Centre for Health Technology Assessment (NCCHTA) that preceded this commissioned work (see below) was whether this type of research could help to identify ways of maximising impact of the programme.

## Main research questions

The agreed proposal set out to address two main questions: the first, a methodological question, required secondary research and the second, an empirical question, required primary research and data collection.

1. *How should the impact of the NHS HTA Programme be measured? What models are available and what are their strengths and weaknesses?*

The principal element of this was a critical literature review of previous work investigating the impact of research programmes, and identifying ways of developing the analysis of key issues. In accordance with the commissioning brief, the aim was not to produce an exhaustive review of all worldwide literature on research impact. Rather, it aimed to provide a comprehensive and insightful qualitative review, which investigated the strengths and weaknesses of the main approaches to assessing research impact, identified models of research impact and provided recommendations for future work in this area. The aim was to identify useful approaches, not to summarise the findings

of all previous evaluations of R&D programmes.

2. *What has been the impact of the first 10 years of the NHS HTA programme from its inception in 1993 to June 2003? What factors seem to be associated with health technology assessment research making an impact?*

Clearly, one aim of the literature review, as specified in the commissioning brief, was to identify how the existing empirical and conceptual literature on research impact assessment could be related to the HTA Programme. The literature findings informed the methods we used to assess impact. However, it was considered not feasible to delay the commencement of the empirical element until the review had been completed. The commissioning brief recognised this and stated that the 'payback approach' from Buxton and colleagues, or an adaptation of it, should form the basis for the assessment. We therefore proposed a multiple methods approach combining quantitative and qualitative methods similar to that used by Hanney, Soper and Buxton in the evaluation of the NHS R&D Implementation Methods Programme.<sup>6,7</sup> Such an approach allows triangulation of methods and data which would include a review of relevant NCCHTA documentation, a questionnaire to lead researchers and a sample of more detailed case studies using interview and documentary analysis.

The aims of this project therefore did not fit neatly into any existing mainstream research. It was not a traditional audit, but neither was it exactly an academic assessment of the performance of the HTA Programme as a whole. The key question set for the project was the second one above, namely an assessment of the impact of the first 10 years of the HTA Programme. It was thought that such a comparatively novel task was best addressed in a two-fold way: through commissioning an exercise using the best known, but still relatively new, payback approach and at the same time undertaking a review of previous work in this field. This dual approach was informed by an exploratory study.

## Exploratory study

The commissioning brief and research proposal were informed by an exploratory study undertaken at NCCHTA in 2002–3. The project, led by Ruairidh Milne, aimed to:



1. clarify the appropriateness, feasibility and resource requirements of different methods of assessing the impact of health technology assessment
2. undertake some health technology assessment impact assessment, as far as resources allowed, providing baseline information about the impact of the health technology assessment for communication purposes
3. make recommendations about how to take forward health technology assessment impact assessment, using either routinely available or specially collected information.

This exploratory project reviewed known key literature, investigated existing information, interviewed R&D 'policy makers' and some project grant holders and developed some case studies. The project confirmed the importance of the issue to the NCCHTA and emphasised the range of different outputs and routes to impact. These included:

- bespoke products for the National Institute for Health and Clinical Excellence (NICE) technology appraisals programme
- products and working relationships in screening [with the National Screening Committee (NSC)]
- the large number of systematic reviews used in NICE guidelines, Clinical Evidence and Cochrane Reviews
- the impact of trials being commissioned and run on communities of practice and specialty groupings
- the high profile of the HTA Programme internationally.

It concluded that although some of the required data were routinely held by the NCCHTA, it was difficult to access and collate. Important but 'softer' information was not recorded at all. It noted the importance of the case studies but that their compilation had proved difficult and slow.

Many of its recommendations focused on the work of the NCCHTA itself, but it also recommended that the HTA Programme should commission research that built on previous work into the impact of R&D and health technology assessment and took a broad view of the possible impacts of health technology assessment. It proposed that a commissioning brief should recommend that the Buxton/Hanney model should be adapted and applied to the HTA Programme over the last 10 years.

## Project team

This research project was undertaken by a team that brought together researchers from the Wessex Institute for Health Research and Development at Southampton University, the home of the NCCHTA, and from the Health Economics Research Group (HERG) at Brunel University. The potential for a perceived conflict of interest is addressed in a separate note and the individual contributions are provided in the Acknowledgements. The research had the benefit of experience and advice from a Project Advisory Group whose membership is also acknowledged.

## The NHS HTA Programme

In 1991, the first formal NHS R&D Strategy was launched. The importance of health technology assessment was recognised in the report on 'Assessing the Effects of Health Technologies' chaired by Sir Iain Chalmers. This led in turn to the establishment of the Standing Group on Health Technology and later the HTA Programme whose aim was to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies was produced in the most effective way for those who use, manage and provide care in the NHS.

A number of key characteristics of the HTA Programme provide an important contextual background to this assessment of its impact. Some of these it shares in common with, but others clearly distinguish it from, many such programmes elsewhere. These characteristics include:

1. As commonly accepted, its definition of health technology covers any method used to promote health, prevent and treat disease and improve rehabilitation or long-term care. 'Technologies' in this context are not confined to new drugs or pieces of sophisticated equipment, but include procedures, settings of care and screening programmes.
2. It has a strong and consistent focus on establishing both the clinical effectiveness and the cost-effectiveness of technologies.
3. It aims to be 'needs-led' and makes extensive efforts to identify and prioritise topics that are important to the NHS as both commissioner and provider of healthcare and to patients themselves.
4. The NCCHTA manages, supports and develops the NHS HTA Programme under contract from

the Department of Health's Research and Development Division. NCCHTA is based in the Wessex Institute for Health R&D at Southampton University.

5. Research proposals are sought mainly through open competition, to address commissioning briefs that characterise the specific needs that have been prioritised.
6. Proposals are peer reviewed by experts and selections and funding decisions made by a Commissioning Board.
7. It funds both secondary and primary research, with projects ranging from brief systematic reviews, some costing as little as £40,000 to large-scale clinical trials costing well over £1 million.
8. Within the Programme there are, or have been, a number of distinctive sub-programmes of work. These include:
  - (a) An important sub-programme of methodological studies. This is now a separate programme and methodological studies were excluded from the remit of this report.
  - (b) Technology Assessment Reports (TARs) commissioned from seven academic research groups under a 'standing' contract. The majority of these are commissioned to provide the TARs for the Technology Appraisal programme of NICE to meet its needs and timescales. These TARs involve secondary reviews of the clinical effectiveness and cost-effectiveness evidence usually combined with some original modelling of cost-effectiveness. (Some studies commissioned from these teams may not be specifically for NICE.)
  - (c) Work is also commissioned to meet specific needs of the NHS National Screening Committee.
9. The HTA Programme publishes its projects in its peer-reviewed monograph series *Health Technology Assessment*. More than 300 'final reports' from HTA projects had been published in this MEDLINE-indexed series by 2005.

These HTA reports were initially available free in paper form, but over time the principal method of access has been as free PDF downloads from the NCCHTA website. An executive summary is prepared for each monograph. Wider publication in academic journals is not precluded, but rather encouraged.

These characteristics make the HTA Programme very different from other health technology assessment agencies where the work is largely carried out by an 'in-house' team, often secondary research without cost-effectiveness and unpublished. Many other health technology assessment programmes operate in 'response mode' to proposals from researchers as opposed to being 'needs-led'.

## Structure of the report

Chapter 2 reports the formal literature review, focusing on methods, but also gives some comparative results from previous empirical studies of other programmes. Chapter 3 summarises the methods used and explains how these compare with previous applications and adaptations of the Buxton/Hanney model. It also highlights the scope and the range of impacts addressed and the element of comparative assessment of questionnaire and case-study findings that was built in to the study. Chapter 4 describes the structure of the HTA Programme, the types of research and the existing monitoring of monographs downloads and of the publications produced. It also describes some evidence about the impact of the HTA Programme that comes from sources beyond our assessment. Chapter 5 details the survey of lead investigators. Chapter 6 presents the results from the 16 more detailed case studies. Chapter 7 draws these elements together in a summary, sets out implications for the HTA Programme and suggests further research on the assessment of the impact of health research programmes.

## Chapter 2

# An assessment of the impact of the NHS HTA Programme: literature review

### Introduction

This chapter presents the findings from a literature review undertaken to identify useful approaches available to assess the impact of health research programmes. The methods used, presented below, are systematic and rigorous. The resulting narrative review synthesises a body of evidence from the existing conceptual and empirical literature on the assessment of the impact of health research programmes.

### Research questions

1. What models are available to assess the impact of health research programmes, and what are their strengths and weaknesses?
2. How should the impact of the NHS HTA Programme be measured?

### Objectives

1. To undertake a critical literature review of previous work investigating the impact of health research programmes.
2. To identify useful approaches for the assessment of the impact of health research programmes.

### Research methods

Relevant literature was identified from a range of sources, including electronic databases, bibliographies of articles, grey literature sources and key experts and organisations.

Identification of the literature/evidence base comprised:

- systematic searching of electronic databases
- project team knowledge in this area
- consultation with experts in this area (circulation of reference listing)
- bibliographic searches of other reviews conducted in this/similar area

- bibliographic searches of references in identified literature.

The search strategy used (see MEDLINE example) provided a pragmatic solution to the problem of searching in a general literature using commonly used text/terms (e.g. health research, assessment, impact). A number of search strategies were piloted and investigated, but there were difficulties limiting the number of 'hits', and the identification of relevant literature (e.g. over 10,000 titles/abstracts from MEDLINE alone). The search terms employed were therefore developed via an iterative process involving an Information Scientist and one of the project team (CG), with feedback during the process from other project team members (SH/CS) (*Box 1*).

### Databases searched

The databases searched were MEDLINE, EMBASE, OVID MEDLINE(R) In-Process & Other Non-Indexed Citations, CINAHL, ECONLIT, Health Technology Assessment Database (HTA), NHS Economic Evaluation Database (NHS EED),

#### **BOX 1** Search strategy (MEDLINE example)

- 1 health services research
- 2 ((health technolog\$ adj3 assessment\$) or hta).mp.
- 3 (NHS adj6 research\$).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
- 4 exp \*Technology Assessment, Biomedical/
- 5 "research and development".mp.
- 6 Health Policy/
- 7 (payback or pay back).ti,ab.
- 8 (("benefit\$" or "utili#ation" or "impact" or "influenc\$" or "gain\$2") adj4 (research or evidence or health technolog\$)).ti,ab.
- 9 ((implement\$ or disseminat\$) adj4 (benefit\$ or impact or gain\$)).ti,ab.

Search terms (1 or 2 or 3 or 4 or 5 or 6) AND (7 or 8 or 9)

Limited to English language and years 1990–2005

HMIC (Health Management Information Consortium), Web of Knowledge (incorporating Science Citation Index and Social Science Citation Index), Cochrane Library, NLM (National Library of Medicine) Gateway Databases and Conference Proceedings Index. Databases were searched from 1990 to July 2005; 1990 was selected as a start date based on (1) knowledge within the team of the literature available, (2) the cut-off date used in a previously published similar methodological review<sup>8</sup> and (3) following consultation with experts in the field.

Bibliographies of identified papers were checked for further references. In particular, previous relevant reviews were examined.<sup>8-13</sup> Experts in the field and key organisations were contacted to identify additional published and unpublished literature, not identified via formal searching methods. The independent Advisory Group for this project was also consulted on a broad listing of key references identified at an early stage of the literature review (Appendices 1 and 2).

## **Inclusion/exclusion criteria (as per project protocol)**

Given the broad perspective of the literature review, the methodology to identify relevant studies comprised an iterative process. This process was guided by the following inclusion and exclusion criteria.

### **Inclusion**

Studies that:

1. describe conceptual or methodological approaches to evaluating the impact of programmes of health research  
OR
2. describe the empirical evaluation of the impact of a particular programme of health research.

### **Exclusion**

Studies that:

1. only provide speculation on the potential impact of proposed (future) research (including recent studies on the value of information)  
OR
2. discuss the impact of research solely in the context of wide and intangible benefits, such as for the good of society and for the overall benefit of the population, without specific measures of impact  
OR

3. only consider impact in terms of guidance implementation.

The above criteria were applied conservatively to the results from the literature search strategy, with studies included in a 'first phase' of the include/exclude process where there was uncertainty. A set of potential references was compiled, including 'probable includes', 'possible includes' and papers of general background interest. This listing of potential key references was then assessed on a case-by-case basis to consider each study against the inclusion/exclusion criteria listed. Inclusion/exclusion criteria were applied independently by two reviewers (SH/CS), with any disagreements resolved through discussion by the two reviewers, and referral to independent assessment by a third reviewer (CG) if necessary and for final decisions.

In the inclusion as studies for this literature review, we were mindful that health technology assessment in the UK has as its main focus the generation of knowledge rather than changing clinical practice. We were also mindful of the fact that internationally many health technology assessment agencies have a role in guidance implementation, and at the outset we clearly stated that it was not our aim to include research which was solely concerned with impact in terms of guidance implementation. The project considered studies that could inform on the assessment of the impact of health research at a 'programme' level, and this was also prominent in the selection of relevant literature. Although there is some ambiguity over what it is that is meant by a 'programme' of research (e.g. a funded programme of research, a collection of related studies/trials, research focused around a specific team), the two-stage include/exclude process was useful in addressing, from a project perspective, whether the key references considered in the second phase of the process were regarded as a 'programme' of research or not. There will be some disagreement generally on the nature of the inclusion/exclusion decisions; however, it is unavoidable in a methodological review of this nature.

## **Literature identified**

Database searching identified about 1600 titles/abstracts. A conservative application of the above criteria together with input from project team members and others resulted in a listing of about 200 potentially relevant references (these are presented in Appendix 1). These references

include studies that are classified ‘includes’, ‘possible includes’ (require more detailed reading to classify) and ‘interest papers’ (that do not meet inclusion criteria).

A more rigorous application of the inclusion/exclusion criteria resulted in a set of studies regarded as a ‘body of evidence’ that met the inclusion/exclusion criteria, when judged on a case-by-case basis. These studies are discussed in the descriptive review (*Table 1*) and are listed in Appendix 2.

## Review methods

The review presents a broad, and summary, descriptive review of the literature identified. Simple data extraction of items such as author details, study methods, setting and characteristics was undertaken using a standard data extraction table (Appendix 3) in order to offer a general summary of a diverse and often complex literature.

Studies are drawn together in a narrative review, with the main frameworks (conceptual approaches) presented in summary format. We thereafter present a brief commentary on the key empirical studies reporting on the application of impact assessment methods and a tabular summary of the findings from the studies most relevant for our study of the impact from the multi-project HTA programme. The discussion includes analysis of the strengths and weaknesses of the conceptual approaches.

## Quality criteria

General texts on quality assessment were consulted (e.g. CRD Report 4<sup>14</sup>); however, we were unable to identify any existing criteria to assess the quality of the conceptual and empirical literature in this area. We used some simple quality assessment questions (see Appendix 3) to investigate the literature in a critical manner. Importantly, these quality assessment questions were not used to identify minimal quality thresholds for the selection of studies or to weight studies in any way. They were only used to guide the interpretation of findings and as an aid in determining a general view on the strengths and weaknesses of the literature.

## Review findings

The review is structured as follows:

- short summary of the literature identified, including in tabular form
- introduction to the research techniques used on impact assessment studies
- presentation of the main conceptual/methodological frameworks identified as suitable for the assessment of the impact of health research programmes
- introduction to the empirical literature on assessment of impact of health research programmes (see *Table 1* for an outline and Appendix 4 for details of studies by framework)
- summary of key findings from those studies that assess the impact of entire multi-project programmes
- discussion including the strengths and weaknesses of the frameworks presented

## Summary of the literature identified

The literature included in the ‘body of evidence’ used to inform the assessment of impact consists of 46 papers. They fall into three categories: five are solely conceptual or methodological approaches, 23 are viewed primarily as applications or empirical evaluations of the impact of a particular programme of health research and 18 are a joint presentation of conceptual approaches and their application. Fourteen of the studies are assessments of the impact of specific health technology assessment programmes. Although a large number of studies are from the UK (14), Canada (10), the USA (7) and Australia (5), the literature identified is of an international nature, with studies from Sweden, The Netherlands, Nigeria, Denmark, Pakistan, Israel and India, plus a multi-national study.

*Table 1* gives a brief summary of each of the 46 papers (for each entry, only one paper is cited; however, in the narrative review multiple references are provided where there have been other related publications), covering basic study characteristics (e.g. type, category, country), the conceptual framework of the study, the techniques applied, the impact categories assessed and the key findings. One column indicates the various actual or potential conflicts of interest that could have arisen in the applications or empirical impact assessments. The main ones are listed here, with the abbreviation that is used to indicate each of them in the conflicts of interest column in *Table 1*:

- Some evaluations are sponsored, or funded, by the body that originally funded the programme of research (sponsor).
- The study team for an assessment is sometimes from the funding body itself (study team).
- The sources of data used for the assessment come from those researchers funded by the research programme, either in the form of questionnaires or interviews (self-report).
- A key person from the centre whose research is being assessed conducts the study themselves (insider account).

The final column in *Table 1* offers some outline comments, including on the relevance of the study, the weaknesses and strengths of each study and the factors associated with impact.

## Techniques used in empirical assessment of the impact of health research programmes

The review identified that four main techniques are used in the empirical assessments of the impact of health research programmes: (1) desk analysis, (2) questionnaires, (3) interviews and (4) case studies; these are used either alone or in combination (e.g. questionnaire followed by desk analysis, interviews plus questionnaires plus desk analysis).

### Desk analysis

Desk (documentary) analysis is a commonly used technique, itself comprising various activities (mainly complementary). These include: documentary analysis of the files about research projects and programmes; analysis of bibliometric databases, particularly those from Thompson ISI (Institute for Scientific Information) that provide information about citations and journal impact factors; and economic evaluation. A particular form of desk analysis, which can result from a combination of desk analysis and the other techniques available (below), is the 'insider account'. Here, one or more researchers describe the impact of research that either they or their colleagues have conducted themselves. This, however, is more likely to be applied to specific studies than to wider programmes.

### Questionnaires

Questionnaires (in various formats) are commonly used as part of impact assessments process. They have most frequently been used to obtain

information from researchers, usually the Principal Investigator on studies but sometimes from all researchers. Often where interviews and/or case studies are to be conducted, a questionnaire will be sent to all projects in a programme so as to gain an overview of the programme. The information can also be used to help identify appropriate interviewees or projects on which case studies can be conducted. Surveys are also sometimes sent to customers and/or potential users of the research.

As with any questionnaires, an important issue becomes the response rate and the attempt to achieve a balance between covering all the issues of interest and restricting the length so as to encourage completion. In terms of the type of questions asked, these can include some combination of closed and open questions. In some instances conceptual frameworks for the assessment studies help inform the structure of the questionnaire and the topics covered.

### Interviews

Interviews can be conducted face-to-face or by telephone and can be with researchers, the customers for the research and/or the potential/actual users of the research. When recruiting interviewees who have used the research, it is often necessary to adopt snowball techniques and ask researchers and other interviewees to suggest the names of those who might have used the research or know of its impact. As with questionnaires, conceptual frameworks for the assessment studies can help inform the structure of the interview schedule. A semi-structured approach is often favoured so as to cover a consistent range of issues in each interview but also allow the particular context and circumstances of each project to be discussed.

### Detailed case studies

Detailed case studies will adopt a range of the above methods, in various combinations. By their nature they are more time consuming and are therefore likely to be restricted to programmes where there are a small number of projects or to a sample of projects. Given the usually uneven distribution of impact between projects in a programme, the selection of case studies is often arranged to ensure that at least some examples of projects with considerable impact are included. Particularly where there has been detailed data collection in a case study, there have been attempts to score the impact from the projects.

TABLE 1 The included studies

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
1	Adamson <sup>15</sup>	1998	APP	Nigeria	College of Medicine of the University of Ibadan: Psychiatry Department	Scientific productivity assessed by desk analysis – bibliometrics; indicators used to show the perceived outcomes of the research studies on the community, but limited details given of how collected	Scientific productivity: considerable. Various career achievements of staff; important impacts on mental health policy and service delivery; helped to get psychiatry accepted as a clinical speciality	Sponsor? Data sources?	Based on one department rather than a programme. Lack of detail on methods used and on link between research and impacts
2	AHFMR <sup>16</sup>	2002	APP	Canada	Alberta Heritage Foundation for Medical Research (AHFMR): 2000–01 HTA products	Interviews (17) with customers for HTA products – open-ended questions	Main outcomes: informed policy, resource allocation, raised awareness	Sponsor	Recognised impact on policy a limited approach, but not resources to go wider; did not attempt to identify impact of specific studies; identified enablers of impact, especially 'committees in place to make required decisions' (i.e. receptor bodies), established communication channels and high quality products
3	AHFMR <sup>17</sup>	2003	APP/ METH	Canada	AHFMR: 2001–02 HTA products	Model of knowledge utilisation that attempted to go beyond instrumental utilisation. Interviews with requestors and/or users and 3 authors (9 in total)	Aimed to go wider than policy impact and consider: knowledge; research targeting; improvements in health services and health. Clients generally satisfied and used products to inform policy	Sponsor	Attempted to move towards a wider multi-dimensional categorisation than limited study reported above, but still insufficient resources. Limitations included possible 'positive response bias'
4	AHFMR <sup>18</sup>	2003	APP	Canada	AHFMR: Health Research Fund projects 1996–2000	Adapted version of the Buxton/Hanney payback approach – questionnaire to PIs, some telephone interviews with decision- makers and users	Impacts included: publications; research training; further knowledge generation; additional qualifications; policy, including guidelines; changed behaviour; health sector benefits	Sponsor; self-report (by researchers)	Provides some endorsement for use of payback framework, but attempted to address identified weaknesses. Very little evidence collected other than that provided by researchers. Research teams with decision-makers or users are more successful than those without

*continued*

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
5	Bjerregaard and Curtis <sup>19</sup>	2004	APP	Denmark (Greenland)	Denmark's National Institute of Public Health: 2 health surveys conducted in Greenland	Assessment of how health survey data used. Presumably desk analysis – insider account, but no account of methods	General description of how the provision of information informed debate, policies, action and further research	Sponsor? Insider account	Not a detailed impact assessment, no account of methods, and not a traditional programme, but uses insider information to illustrate impact of public health research
6	Bos <i>et al.</i> <sup>20</sup>	1997	APP	Multi- national – 5 countries	HTA reports on invasive cardiology therapy for CHD in 5 countries	Questionnaires, database information, publications, review of draft papers	Identified effect of HTA studies on coverage decisions. Most HTAs in this field conducted by official bodies have had some effect, but decisions based on many issues	Study team included the experts from each country, but also others	Not really based on a full programme. But some triangulation of methods. The availability of a structure for HTA an important prerequisite for those HTA studies to have any impact on coverage decisions
7	Britton and Jonsson <sup>21</sup>	2002	APP	Sweden	Swedish Council on Technology Assessment in Health Care (SBU): 7 SBU reports (systematic reviews)	Desk analysis: monitor changes in practice after the systematic reviews using secondary analysis of survey data collected by others, but specific methods not explicitly stated	Examined the changes in medical practice and suggest the reports had at least some influence in the identified changes	Sponsor? Study team	Not based on a full programme. Limited discussion of methods which appeared to be restricted to the secondary analysis of cause and effect, i.e. would changes have occurred without reports? Admitted difficult to pinpoint influence of SBU reports
8	Brownson and Simoes <sup>22</sup>	1999	APP/ METH	USA	Centre for Disease Control and Prevention: prevention research – Environmental tobacco smoke (ETS)	Framework to consider process measures, impact and outcomes. Case study – secondary analysis of existing materials – linked to the causal model and 5 areas of public health, but methods not clearly presented	Impact indicators include changes in knowledge, attitudes or risk factor prevalence; outcome evaluation measures changes in morbidity, mortality, and quality of life, also cost- effectiveness	Sponsor	It is not clear how far the focus is on research or health preventions, and even where it is on research it is not on a specific programme. No rationale provided for selection of case study. The methods are not really clear, beyond a secondary analysis. But a comprehensive list of impacts and outcomes

continued



TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
9	Buxton and Hanney <sup>23</sup>	1996	METH	UK	General review	Paper presents the payback approach in detail: framework of multi-dimensional categorisation of benefits and model of how to assess them. Presents 8 case-study projects/areas for illustration of the payback approach	5 main categories in the multi-dimensional categorisation: knowledge production; research targeting and capacity building; informing policy making; health and health sector benefits; wider economic benefits	Sponsor	Detailed presentation of the payback method (for the first time), but application in case studies not to a programme. Factors identified as being associated with high levels of payback in some of the case studies included: continuing support from customers and active brokerage; liaison with stakeholders; appropriateness and quality of the research and active dissemination; working in a research unit
10	Buxton et al. <sup>24</sup>	1999	APP/ METH	UK	NHS North Thames Region: R&D Programme	Used payback framework. Used questionnaires and some bibliometric analysis for all projects and case studies (19). Benefit scoring system used to score questionnaire responses and re-score the case studies	Multi-dimensional categorisation used. From questionnaire some attempt to quantify knowledge production, contribution to future research and research training, and claims for impact on policy and health gain. More details in case studies. Limited examples of wider economic benefit	Sponsor; self-report (by researcher)	Demonstrated application of payback framework to a full research programme. Difficult to apply later stages of payback categories but made recommendations about future assessments including a revised questionnaire and informing researchers from the start of possible impact assessment demands. No clear relationship established between score for impact on policy and practice and bibliometric indicators. Considerable impact from a funded research centre
11	Buxton and Schneider <sup>25</sup>	1999	APP/ METH	Canada	AHFMR: 2 external projects/2 internal HTA projects/4 programmes of research	Payback methods used. Desk analysis, interviews with key researchers. No questionnaire used. Sequence of research seems to be main focus – not dimensions of payback.	Impact not presented in any detail. Payback dimensions not used to categorise payback from programmes/projects	Sponsor; main study team researcher; self-report (by researchers)	Findings mainly methodological, i.e. found payback method worked and was useful. Able to generalise the approach from UK to Canada. Limited to a 2-month period of activity. Conclusions not fully supported by the data presented. Speculative discussion of payback, rather than actual evidence of payback

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
12	COHRED <sup>26</sup>	2000	APP	Pakistan	3 child health programmes (not specifically research programmes, but included some research)	In-depth interviews with 16 policy makers, and 22 researchers. Review of research studies and programme documents. Looked at impact of any relevant research on policy and programmes	Research little used overall in policy making. Gap between research and policy	Some self-reporting	Impact of specific programmes of research not really addressed
13	Croxson et al. <sup>27</sup>	2001	METH	UK		Presents the payback approach as a basis for a system of routine performance of an R&D programme, taking UK NHS R&D as a possible example		NA	Methods paper
14	Dixon et al. <sup>28</sup>	2003	APP	UK	NHS South and West Region: Development and Evaluation Committee's (DEC) technology appraisal reports	Multifaceted study design used surveys of clinicians and managers, interviews with selected respondents (30). Asked about the awareness and use of 7 sources of evidence based guidance without indicating focus of the study. Desk analysis of trends in utilisation before and after publication of 4 reports used as case studies. Compared impact in the S&W region and in a control NHS region and random health authorities	Examined impact on policy decisions and clinical practice. Considerable impact in S&W but not elsewhere and impact on informing policy but actual impact on practice could not be identified in routine data	Sponsor	The multifaceted study design is important, but the DEC reports are more like NICE guidance than HTA reports therefore this is perhaps assessing the impact of appraisal more than HTA research? Authors recommend its use for evaluating impact of NICE. Respondents in the local health region claim that the quality of the information was more important than any other factor in influencing their considerable use of the reports, yet use much higher locally than elsewhere. This raises important questions about research location

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
15	Drummond et al. <sup>29</sup>	1992	APP/ METH	USA	National Eye Institute: major clinical trial for early detection of diabetic retinopathy	Economic analysis/CBA using a decision model, expert opinion and cost data to make an estimate of expected benefit from the completed trial. Decision tree model of disease and treatment. Sensitivity analysis	Looked at outcomes (disability days avoided/QALYs gained, and cost savings) and showed expected impact in terms of economic and health benefits from the trial. The economic benefits are in the form of the human capital approach of considering the benefits of the production loss avoided	Sponsor	Not looking at impact of research programmes, looks at possible impacts (decision modelling). Method is a decision model – is this a means of assessing impact of a programme? Retrospective assessment of potential net social benefit? Method also presented as way of looking at proposed research and overall an important early exploration of methods
16	Eisenberg <sup>30</sup>	2001	APP/ METH	USA	Agency for Health Research and Quality (AHRQ): study on one study – Consumer Assessment of Health Plans	Describes framework previously developed <sup>34</sup> consisting of a pyramid of outcomes with 4 levels. A case study results presented, but no methods stated – presumably secondary data?	The 4 levels of outcomes in the pyramid start at the base with knowledge and further research; then impact on policies; practice; with health outcomes at the apex (has similarities with payback). Impact at all levels claimed within the case study but no demonstration that the impact on people's decision making has led to an impact on health outcomes	Sponsor; study conducted by Director of AHRQ	No methods stated. No rationale for case study used? Is primarily a commentary on the importance, to secure future funding, of health research bodies demonstrating to politicians the benefits that come from the research funding. Does repeat an important model of impacts

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
17	Elliott and Popay <sup>31</sup>	2000	APP	UK	NHS North West Region: 9 case studies of purchaser initiated research	Used a model of research utilisation in policy making based on the problem solving and interactive models of Weiss and the dialogical model of Giddens – social knowledge is jointly constructed from the interaction between researchers and others. This is a model of research utilisation rather than impact assessment. Case studies consisted of interviews and desk analysis of documentary evidence, and testing validity of findings on participants	Examined how local NHS policy makers used research. Found mostly indirect impact, e.g. in policy debate, but few attempts to describe the impact of specific studies	Sponsor	Not really a full a programme. Limited details about methods of impact assessment and there are problems in operationalising the dialogical approach but important analysis of how sustained dialogue improves the utilisation of research
18	Ferguson <i>et al.</i> <sup>32</sup>	2000	APP	UK	NHS Northern and Yorkshire Region: 3 broad programmes – biomedical, HSR, primary and community care (158 projects)	Refer to payback framework and do not attempt to develop own framework. Used questionnaires to all researchers to gather quantitative and qualitative data, and desk analysis (bibliometrics)	Examined publications, presentations, changes in individual practice, changes in NHS service delivery, career development of researchers. Reported impact on publications, more effective treatment, benefits to commissioning, research capacity, other research funding	Sponsor: most of study team; self-report (by researchers)	Study not strongly related to a conceptual framework. Data gathered on a wide range of impacts and authors argue success rates appear impressive. While data collected on presentations and non-peer-reviewed output, they did not add to the descriptive statistical analysis. Difficulties in collating a reliable database of publications usefully illustrate technical difficulties in impact studies

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
19	Hailey et al. <sup>33</sup>	1990	APP/ METH	Australia	National Health Technology Advisory Panel (NHTAP) reports	Method: presents 6 measures of impact (but not techniques re operational approach). APP – looks at technologies (20) covered by HTA reports up to end of 1988, reports from NHTAP. Little provided on methods – presumably desk analysis: just states comparing recommendations, assessments, policy activities. Examines requests received for copies of the reports, citations of reports	Looked at impact on policy as main factor; but also impact as educational material, and practice. Out of the first 20 technologies covered by HTA reports there had been significant impact in 11 and probable influence in 3	?	At the margin of our inclusion on criteria because based more on NHTAP recommendations than a programme of research. Absence of methods information. But early paper – important to stress – and describes complexity of impact assessment, with a range of dimensions of impact. Timing crucial factor for impact
20	Hailey and Crowe <sup>34</sup>	1993	APP	Australia	NHTAP: MRI (diffusion of) following a multi-site HTA and NHTAP synthesis report	Appears to be desk analysis: looks at data on MRI from health authorities, radiology departments, distributors of MRI equipment and the Health Insurance Commission	MRI diffusion shown, with impact on policies, funding and use	?	Is not a full programme, but a case study based on a multi-site trial. Limited details on methods but provides a useful example of impact

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
21	Hailey <sup>35</sup>	1993	APP	Australia	Australian Institute of Health and Welfare (AIHW): 45 HTA reports (some of which were guidelines)	Refers to an approach by Drummond, Hailey and Selby Smith that considers the relationship between health policy, administration and technology. Claims the impact of an assessment can be inferred through the outcomes of major recommendations made, citations, requests for report, and subsequent deployment and use of each technology; but no details given about the specific techniques used in this paper	Look at impact on health policy. Impact on policy identified for 17 of 26 technologies assessed. The majority of major recommendations in the reports accepted when they related to introduction of the technology, reimbursement or further research	?	The paper may not be about research programmes? Lack of details on methods and difficult to judge how the impact assessed on this basis? Important analysis of circumstances in which HTAs have an impact: "The conditions under which the advisory bodies undertook their work need to be considered in any discussion of the impact of their evaluations"
22	Hailey et al. <sup>36</sup>	2000	APP	Canada	Canadian province (not stated): 20 short HTA technical notes – specifically requested by decision- makers	Interviews with those requesting the HTA notes – covered clarity, relevance of advice, whether helpful. Checks on quality of the reports made using desk analysis (subsequent HTA reports, monitoring literature) and comments from experts	Impact on policy and other decisions considered: 14 had influence on policy decisions	Sponsor?	Perhaps the focus is more about the product (i.e. are Technotes useful HTA?) than about the research undertaken that is discussed by the HTA technical note? Some lack of detail about the impact method used. Did show the usefulness of Technotes within the HTA programme and that they were making an impact

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
23	Hailey <sup>37</sup>	2003	METH	Canada		Examines the use of organisational theory approach to measuring effectiveness of HTA. Assessment of impact part of a wider evaluation of an HTA programme. Suggests survey approaches – or more elaborate studies. Examination of correspondence, documents, publications, administrative data	Impact items to include: whether report considered; recommendations accepted and material from an HTA report is incorporated into policy documents, referenced in guidelines, etc.	NA	Paper makes an important contribution in considering the role of impact assessment within a wider evaluation of an HTA programme and claims the effectiveness of an HTA programme is determined by many factors including downstream components that are largely not under the control of the programme. Impact assessment difficult, therefore evaluation should cover: "activity, continuity, relevance, quality and impact"
24	Hall <sup>38</sup>	2004	APP	Australia	National Drug and Alcohol Research Centre (NDARC): research – core funded by NSW government	Mainly 'insider accounts' – discussions with colleagues – and desk analysis – analysis of documents/database	Looks at impact on policy formulation – also on further research. Claims such impact but basis for saying so seems a bit loose	Insider account	Lacks a framework for impact assessment. General comments useful

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
25	Hanney et al. <sup>6</sup>	2003	APP/ METH	UK	NHS: Implementation Methods Programme (36 projects)	Payback framework used, but impact not the main focus of this evaluation. Questionnaires to lead researchers (postal), to users (electronic and postal). 83% response to questionnaire from lead researchers. Poor response from users to electronic survey. Some desk analysis: bibliometrics. Interviews with key figures in programme for evaluation of commissioning, etc.	For impact assessment concentrated on: publications; research targeting and training, informing policy and practice. User surveys and interviews provided a limited amount of verification of the impact claimed by researchers' self-reports	Sponsor: Data on most projects was self-report (by researchers)	Nature of impact assessment not always clear within wider evaluation and impact of overall programme not very clear. Problem of using cross- sectional approach to assess impact of a programme highlighted by this study. High response rate achieved. Useful in providing a prototype for Table 2 used in the current report
26	Jacob and Battista <sup>39</sup>	1993	APP/ METH	Canada	Quebec Council on Health Care Technology Assessments (CETS): first 4 years' work (10 reports)	Case study analyses of impact on decision- making and cost savings. Reports scored for policy influence – critical incidents used. Interviews (45) with scientific and political partners and staff at CETS. Documentary analysis also used. Desk analysis of cost savings. 2 case studies described in detail	Examined impact on decision- making and cost savings. 8 of 10 reports influenced decisions	Sponsor; study team lead; some data sources	Precursor to later larger study by Jacob and McGregor. <sup>40</sup> Identified a series of key features of the Quebec system that were favourable to HTAs making an impact. These include: "a general receptivity to rationality in decision making ... the healthcare system in Quebec is organised in such a way that information produced by the council can filter easily into the decision- making process"

continued



TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
27	Jacob and McGregor <sup>40</sup>	1997	APP/ METH	Canada	Canada/ Quebec Council on Health Care Technology Assessments (CETS): first 21 reports	Some mapping of the HTA process – followed by case study analyses. Similar to above. Comprehensive case study approach	Looked at changes in policy, and changes in distribution of the technology, e.g. policy change via guidelines of professional bodies (for decision by doctors or hospitals). When government decision impact on policy more difficult to establish. Also looked at impact on direct costs on health care system. 18 of 21 reports had influence, 8 at highest level	Sponsor; study team lead; some data sources	In limitations the authors say methods heavily dependent on judgement of the analyst. Therefore, confirmation of impact by key actors is important. Detailed discussion of comprehensive methods, including scaling or scoring of projects but this not used to compare average impact of different types of report. A key paper in development of HTA impact assessment
28	Jonsson et al. <sup>41</sup>	2001	APP	Sweden	SBU: SBU reports on 3 screening procedures (+ literature)	Used reports and literature to identify impact. Insider account?	As part of wider Eurassess project looked at impact on policy and practice. Found HTA an important part of policy making in Sweden	Sponsor as part of Eurassess project? Study team	As part of Eurassess project looking at just 3 screening procedures, not a full programme. Could be more of a commentary paper than an impact assessment study
29	Kerridge et al. <sup>42</sup> (abstract only)	2004	APP/ METH	UK	NHS HTA Programme	Pilot study for the current study. Interviews, literature, database. Looked at case studies	Impacts include: tailored products for NICE, products in the area of screening, and the impact of trials on communities of practice and specialty groups	Sponsor; study team	General pilot study and only described in abstract. More a case of looking at how impact may be assessed. Suggests Buxton/Hanney payback as useful approach
30	Lavis et al. <sup>43</sup>	2003	METH	Canada		Mapping of research process – and target audience important. General paper presenting a range of approaches/techniques that could be used – files, surveys, interviews	Different audiences warrant different measures of impact, but the use of research knowledge to inform decision- making, not a change in health status, constitutes the most appropriate generic measure of the impact of research that can be expressed routinely	NA	Very comprehensive analysis of techniques, etc. but may not be best suited to assessing impact of programmes? Suggests focusing on impact on policy not health status

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
31	McGregor and Brophy <sup>44</sup>	2005	APP	Canada	McGill University Health Centre, Quebec: HTA Unit	Presumably informed by previous studies in Quebec (see above) but no account of concepts or techniques. Presumably desk analysis of documents etc and insider account	Examined impacts on hospital policy towards the therapies subject to the HTA reports and cost implications. All reports incorporated into hospital policy and some cost savings	Sponsor; study team	More of a commentary on the work of HTA Unit than a separate assessment: no details of assessment techniques. HTA Unit combines researchers to synthesise the evidence and a policy committee to make recommendations. Provides useful insights at a hospital system level
32	Molas-Gallart et al. <sup>45</sup>	2000	APP/ METH	UK	ESRC: AIDS programme – on social aspects of AIDS (15 projects)	Framework based on the interconnection of 3 major elements: the type of output that can be expected from research; the channels through which their diffusion to non- academic actors occurs; the forms of impact. Interviews to map network of researchers and users. Tracing of post- research activity. Interview with researchers – then snowball approach for user interviews. For application 13 sets of interviews (researchers and users): 43 interviews	Types of impact: use of skills and tacit knowledge to solve specific technical problems; use of codified knowledge in the form of models to solve problems; influence decisions either policy or technical including confirmatory or justification. Programme did provide non-academics with tools to solve problems and been used to develop policies	Sponsor	How far the full comprehensive framework was helpful is not clear, but it does demonstrate the complexity of assessing research impact. Many useful lessons provided by study: concludes a 2 to 3 stage process required to assess impact (interview researchers first, then users); normal sampling techniques inadequate because impact not distributed along a normal distribution curve; considerable diversity of communication and therefore, “aggregate data on the percentage of projects that have established stable links with users are of limited value”; detailed project-by- project qualitative analysis important; wait 1 to 2 years to assess

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
33	NIH <sup>46</sup>	1993	APP/ METH	USA	National Institutes of Health: 34 examples of NIH-funded research	Conceptual approach is to produce a monetary figure for cost savings, includes both direct and indirect cost savings, i.e. uses human capital approach with consideration of lost earnings avoided. Method used is speculative, assume a rate of adoption for a technology based on expert opinion to produce prospective savings from a cohort (using desk analysis)	The benefit is expressed as a monetary cost saving. Estimated monetary benefits greater than research costs in the cases conducted, but for each case a high and low estimate is given	Sponsor; study teams; some sources of data	Most of the cases involve estimation of probable impact rather than conducting actual studies, not research on evidence of impact. It is also a few highly selective studies from the full portfolio of NIH work. There are now reservations about using the human capital approach and studies suggest the NIH work might not be the sole cause of benefits (Buxton <i>et al.</i> <sup>10</sup> ). But these are very detailed studies and do at least show the considerable potential benefits from health research
34	NHSE South West <sup>47</sup>	1999	APP	UK	NHS South West Region: 162 completed projects	Used end of project assessment form used to gather data on publications/dissemin- ation/higher degrees/other outputs/further research. Desk analysis: bibliometrics	Impacts examined limited to publications, research training and targeting. 55 projects reported publications/39 led to higher degrees/73 reported active dissemination	Sponsor; study team; self-report (by researchers)	Much of the focus is on processes and limited range of impacts analysed. But 109 questionnaires returned (74%) and study informed the introduction of more formal assessment of the potential impact and benefit to NHS of completed projects
35	NHSE Trent <sup>48</sup>	1997	APP	UK	NHS Trent Region: Responsive funding scheme and HSR Training Awards	Used a version of Payback methodology. Questionnaire/survey and some interviews. 80–90% response to survey	Looked at outputs, research training, future research, impact on policy and practice. Found high level of outputs but less impact on policy	Sponsor; study team; self-report (by researchers)	Study was seen as only a first step in the evaluation of payback, but did provide evidence to inform policy on research programmes

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
36	Raiten and Berman <sup>49</sup>	1993	APP/ METH	USA	Monoclonal antibody research (history of the research in this area rather than a specific programme); benefits from one application – HIV testing of blood supply	Cost-benefit analysis (CBA) to put monetary figure on impacts from basic biomedical research. Economic evaluation approach. Historical tracing of research that led to key findings on monoclonal antibodies. Then conducted a case study on benefits of just one of the many applications	Benefit: measure in income losses avoided, reduce medical care costs, production outputs. A substantial return on investment is demonstrated, but benefits in USA and costs of research fell on UK. Benefit to cost ratio at 19:1	Not direct, but indirectly sponsoring organisation keen to show benefits of research	Examines a stream of research rather than a specific programme. Some difficulties with the methods used to establish the impact, for example it is not clear how the historical tracing was conducted or how rate of return calculated when costs were to UK and benefits to USA. Does attempt to address criticisms of previous micro CBA for focusing narrowly on applied research and macro CBA for being aggregative and inconclusive
37	Royal Netherlands AAS <sup>50</sup>	2002	METH	The Netherlands	Sets out the societal impacts of health research and a range of output categories and possible impact indicators to identify them: content analysis/citations/auth- orships/products/ funding of research/ memberships/ teaching, etc. Not really presented as a conceptual framework but a way of applying it described: through self evaluation and then judged by evaluators using stakeholder comments	The societal impacts are: relevance for healthcare providers and the process of healthcare delivery; relevance for policy makers and the process of designing, implementing and monitoring policy decisions. Health gain seen as a final objective but a challenge to evaluate requiring extensive research: use the indirect indicators instead	NA	Perhaps rather lacking in an overall conceptual framework and this possibly contributes to some lack of distinction between processes and output. Does though provide a large range of criteria and matching indicators and sets out how they should be assessed: self- evaluation by researchers and then evaluators judge it using stakeholder comments. If it is operationalised this would provide a way of bringing impacts into the regular assessment of research teams. Even then it is more relevant for the work of teams rather than programmes	

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
38	Shah and Ward <sup>51</sup>	2001	APP	Australia	National Health and Medical Research Council (NHMRC): all grants awarded in 1993 (32 new, plus 31 continuing) by the Public Health Research and Development Committee	Self-complete questionnaire completed by chief investigators (telephone call to investigators prior to despatch of questionnaire). Combined with desk analysis: database information on journal impact factor, and listing of publications in database (MEDLINE/EMBASE). Attempted some correlations between publications and impact on policy and practice	Questionnaires covered a broad range of impacts: publications, presentations, research training, influence on policy and practice. Findings: 82% produced publications/research presentations/58% stated research influenced policy/69% stated influence on practice/53% stated both	Sponsor not stated; self- report (by researchers)	Authors acknowledge some methodological limitations such as looking at the programme's grants for 1 year only, but do not comment on almost total reliance on self-completed questionnaires. Do include a wide range of impacts, report positive findings, and state, "We offer our survey instrument as a prototype tool" (p. 560)
39	Shani <i>et al.</i> <sup>52</sup>	2000	APP	Israel	Medical Technologies Administration (MTA) of the Israeli Ministry of Health: all 84 (?) 1999 technology reports	Commentary on the process and approaches used by the MTA in assessing a new list of technologies in 1999 for possible inclusion in the National List of Health Services – no methods stated for the commentary. Insider account?	Examined the influence of the MTA on the coverage policies made. Most of the recommendations were accepted by policy makers	Sponsor?	Again the body in question is more of an appraisal agency than a research funder, but does undertake HTAs. No methods stated for this study and no clear evidence base. It is though included in the Gerhardus review of impacts of HTAs and is an example of where the institutional arrangements are established for HTAs to make an impact

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
40	Spaapen and Sy/vain <sup>53</sup>	1994	APP/ METH	The Netherlands	Dutch Sector Council: pilot on 2 health research centres	Complex model/framework through which research (not specifically health research) can be assessed with respect to its potential value for society. Discuss general process issues. Survey/questionnaire. Interviews. Bibliometrics. Expert opinion. Uses "societal quality research profiles" to present findings	The framework has 3 main assessment categories: (1) different forms of research output (written and non- written); (2) different forms of societal demand (concrete and strategic); (3) different media through which (1) and (2) interact. Illustrated in 2 case studies	Some self- reporting?	Primarily a methods paper, setting out a methodological framework that is quite ambitious and goes wider than health, but tested on 2 health research centres. More relevant for centres than programmes, but dimensions of societal value covered are rather modest. The societal quality research profiles are a useful tool and pre-date the equivalent HERG/RAND payback profiles (see Wooding et al. <sup>5</sup> )
41	Stryer et al. <sup>54</sup>	2000	APP/ METH	USA	AHRQ: all outcomes and effectiveness research from 1989 to 1997 (91 PIs)	Uses a model (based on desk analysis) to consider data already collected in questionnaire/survey of PIs. Presented in the form of a triangle with 4 levels: the widest at the base, level 1, is research findings, going through to level 4 at the apex: impact on healthcare outcomes	Looks at impact in terms of research findings/publications/impact on policy/impact on practice/impact on outcomes. Found limited impact on policy, practice and outcomes	Sponsor; self-report (by researchers)	Despite fairly broad and detailed conceptual framework, authors observe it does not capture some important more general contributions of AHRQ's OER, e.g. it has played a major role in the expansion of interest and capacity in outcomes research. The background to the study was the pressures created by the GPRA. The items in the levels are similar to the payback multi-dimensional categorisation

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
42	Weisbrod <sup>55</sup>	1971	APP/ METH	USA	Programme of work by Stalk and Sabin that led to vaccines for polio	CBA – benefits worked out in monetary terms. Uses economic estimates of impact and values the benefits by a human capital approach which considers the market value of production lost by mortality and morbidity caused by polio as well as the costs of treatment and rehabilitation of victims. Various alternative assumptions are made. (Presumably desk analysis)	Impact stated in terms of a rate of return on the research and application costs. In this case at least 5% and more likely 11–12%	None	Uses the human capital approach to provide the benefits data for the CBA and rate of return and this approach is now subject to more criticism. Seen as early/classic study, but there have not been the follow-on studies Weisbrod hoped for. Perhaps this was a particularly clear example on which to conduct the study but other topics more difficult
43	Wisely <sup>56</sup>	2001	APP	UK	NHS: National R&D programme on primary/ secondary care interface (70 projects – 64 completed by time of assessment)	Based on payback approach. Questionnaire to lead researchers. Plus use of programme records (desk analysis). Also looked at cost of research and overspend, delays, etc. Correlations made between response rate and projects that were more highly rated in the initial review	Used multi-dimensional categorisation: publications, higher degrees, research targeting, impact on policy, practice and health. Impact was reported in all these categories, though more expected than already occurred. More detailed accounts given in some questionnaires	Sponsor; study team; self-report (by researchers)	Researchers' recommendations suggest use of case studies

continued

TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
44	Wisely <sup>57</sup>	2001	APP	UK	NHS: National R&D programme – Mother and Child Care (51 projects in 21 areas – 39 completed by time of assessment)	Based on payback approach. Questionnaire to lead researchers. Plus use of programme records (desk analysis). Also looked at cost of research and overspend, delays, etc. Correlations made between response rate and projects that were more highly rated in the initial review	Used multi-dimensional categorisation: publications, higher degrees, research targeting, impact on policy, practice and health. Impact was reported in all these categories, though more expected than already occurred. More detailed accounts given in some questionnaires	Sponsor; study team; self-report (by researchers)	Author suggests researchers might not be best placed to judge the impact of research findings
45	Wooding et al. <sup>5</sup>	2004	APP/ METH	UK	Arthritis Research Campaign (ARC): 16 research grants awarded in early 1990s	Payback approach. Looked at 16 research grants – through case studies of semi-structured interviews with key informants, documentary analysis, bibliometric indicators. Findings were scored by research team and the results presented as payback profiles, or spidergrams. This enabled findings from different groups (types of research, mode of funding, etc.)	Used multi-dimensional categorisation: publications; higher degrees and research targeting; impact on policy; health and health sector gains; wider economic benefits. Impacts found in all categories, but more in the early categories with just 2 of 16 showing economic benefits, 3 speculated economic benefits	Sponsor; most interviews were with researchers	The case studies were carefully selected using bibliometric analysis and expert consultation, but only covered 16 out of 556 awards between 1990 and 1994. It was difficult to assess whether no evidence meant there might be impact but there was no evidence or meant there was no impact. Limited economic benefits. The detailed analysis and findings showed there is a diversity of research payback, short focused project grants seem to provide value for money and the payback framework could be operationalised and embedded by ARC

continued



TABLE 1 The included studies (cont'd)

No.	Authors	Year	Type	Country	Programme/ specialty	Concepts and techniques/methods	Impact: examined and found	Conflicts of interest	Comments: meeting inclusion criteria; strengths and weaknesses; factors associated with impact, etc.
46	Yesudian <sup>58</sup>	2001	APP	India	Research by 2 centres on the regulation of the private health sector in Bombay	Case study based on desk analysis: insider account. More of a commentary	Examined impact on policy. Research findings said to facilitate the policy process through dialogue based on reliable evidence	Sponsor? Partial insider account	Is looking more at a centre than a programme and is more of a commentary. Provided insights and author claims study "shows the importance not only of research findings but of a trusted institutionalised research process that enables ... continuous policy dialogue"

APP, applications; CBA, cost-benefit analysis; CHD, coronary heart disease; ESRC, Economic and Social Research Council; GPRA, Government Performance and Results Act; HSR, health services research; HTA, health technology assessment; METH, methods; MRI, magnetic resonance imaging; NA, not applicable; OER, Outcomes and Effectiveness Research; PI, Principal Investigator; QALY, quality-adjusted life-year.

## Conceptual or methodological approaches identified for the assessment of the impact of health research programmes

The review has identified the following approaches for the assessment of the impact of health research programmes:

- the ‘Payback’ approach (e.g. reference 23)
- monetary value approach to estimating returns from research (i.e. cost–benefit analysis, or estimated cost savings)
- the framework developed in the Economic and Social Research Council (ESRC) project on the non-academic impact of socio-economic research<sup>45</sup>
- the framework developed in the study of the Royal Netherlands Academy of Arts and Science (2002) to assess the societal impact of applied health research (partly informed by methods from Spaapen and Sylvain<sup>53</sup>)
- framework presented by Lavis and colleagues (Canada) (e.g. reference 43), to assess the impact of health research on policy (with a focus on a specific target audience)
- organisational theory approach/framework presented by Hailey (e.g. reference 59)
- detailed case studies and follow-up analysis, on health technology assessment policy impacts and cost savings in Quebec, Canada (e.g. reference 40)
- the Knowledge Utilisation model developed in Alberta, Canada, around research outputs of AHFMR (e.g. reference 16)
- Impact model developed by the USA Agency for Healthcare and Research Quality (AHRQ) (e.g. references 30 and 54).
- plus other non-specific impact projects, outside of the above categorisation (see Appendix 2).

We present some detail on each of the above frameworks below. The detail presented varies according to the level of detail identified in the literature search and on the perceived suitability of the approach for the assessment of the impact of health research programmes (such as the NHS HTA Programme). Interested readers are urged to consult the references cited for further detail.

### Buxton and Hanney ‘payback’ approach

Buxton and Hanney<sup>23</sup> presented the ‘payback’ approach as a framework for the assessment of the impact from health research. The approach considers five categories of ‘payback’ from health services research (Box 2). It is presented as an

‘input–output model’, but also captures many of the characteristics of earlier models of research utilisation (e.g. references 60–62). Importantly, the payback approach makes a clear break from the terminology of ‘utilisation’ and/or ‘use’ of research and provides a comprehensive framework for assessment of impact, i.e. the concept of payback.

The payback approach offers (1) a framework for describing the sequencing of the research process, from needs assessment to dissemination, and (2) a multidimensional categorisation of benefits, a range of potential benefits (payback), from health services research. Both of these contributions have been helpful in the empirical work undertaken, using the payback approach, to assess the impact of health services research (e.g. references 6, 24, 25).

Knowledge is the first step in the payback approach, and this reflects the published outputs from research (e.g. journal articles), but may not be limited to published materials alone (i.e. other methods of dissemination). The other categories of payback capture the broad array of possible payback, and measurement of impact, from health services research. Category 2 reflects the broader benefits to future research and research use, and

#### BOX 2 Categories of payback (from Buxton and colleagues<sup>24</sup>)

- 1. Knowledge**
- 2. Benefits to future research and research use:**
  - (a) the better targeting of future research
  - (b) the development of research skills, personnel and overall research capacity
  - (c) a critical capability to utilise appropriately existing research including that from overseas
  - (d) Staff development/educational benefits
- 3. Political and administrative benefits:**
  - (a) improved information bases on which to take political and executive decisions
  - (b) other political benefits from undertaking research
- 4. Health sector benefits:**
  - (a) cost reduction in the delivery of existing services
  - (b) qualitative improvements in the process of service delivery
  - (c) increased effectiveness of services, e.g. increased health
  - (d) equity, e.g. improved allocation of resources at an area level, better targeting and accessibility
  - (e) revenues gained from intellectual property rights
  - (f) organisational development
- 5. Broader economic benefits:**
  - (a) wider economic benefits from commercial exploitation of innovations arising from R&D
  - (b) economic benefits from a healthy workforce and reduction in working days lost.

although this may have some overlap with aspects of the knowledge generated, it captures other important benefits from research, i.e. use of knowledge to target further research funding, development of research capacity/skills (either through specific training or via improvements in the ability to use existing research). Category 3, political and administrative benefits, seeks to inform on the payback that research has provided in the area of policy decisions (debates), i.e. to inform and/or justify decisions. Category 4 seeks to measure the direct impact of research on the healthcare system, via cost impact (e.g. cost savings), health service delivery, use of more effective health services (leading to improvements in health) and on the distribution of healthcare (e.g. amongst different patient groups or geographical regions). Category 5 sets out the broader economic benefits that could be attributed to payback from health services research.

The research sequence is mapped out over seven stages, with interfaces between (1) the research system and (2) the wider health system. The process allows for consideration of measurement of payback at each of the stages and interface activities.

The payback approach is a useful conceptual approach to the measurement of impact from health services research. However, importantly, the authors present it as a practical tool for the assessment of payback, demonstrating (through case studies) the potential for the payback approach to be used to guide the empirical assessment (quantification) of payback from specific research programmes and activities. Buxton and Hanney<sup>23</sup> present eight simple case studies (in outline) to demonstrate the usefulness of the approach. This initial demonstration of payback methods does not provide guidance on the research techniques used to operationalise the approach. However, over the past decade a number of large-scale studies have been completed using the payback approach. There have been variations in the way the practical assessment of payback has been undertaken, with studies employing a range of techniques (e.g. questionnaires, interviews, documentary analysis, detailed case studies), either alone or in combination. Although there have been variations in the way the practical assessment of payback has been undertaken, there has been a common adoption of the payback framework to guide these 'payback' studies (see detail below). Indeed, the payback approach is not presented as a 'one size fits all' approach, but as a tool that can be applied

to suit a particular research setting. Buxton and Hanney highlight the limitations of the approach, indicating that it would be unreasonable to expect most projects to produce payback in all categories. They suggest the approach could be useful when applied in a context specific way when assessing research centres. Efforts have also been made to show how the framework could be adapted for use in regular monitoring.<sup>27</sup>

### **Monetary value approach to estimating returns from research (cost-benefit analysis or estimated cost savings)**

A number of studies<sup>29,46,49,55</sup> have attempted to value the impact of specific health research in terms of a single monetary value. This has taken the form of analysis of cost savings or a 'cost-benefit analysis' (CBA). The precise nature of these studies and the items taken into consideration has varied. In the cost savings approaches, both direct cost savings (e.g. cost to the health system) and indirect cost savings (e.g. costs to the patient or community) can be estimated. Direct cost savings (or reduction on claims on resources) come about as a result of health research that leads to new treatments that reduce the overall cost per patient or the number of patients that need to be treated. Such savings may also accrue more widely than just to the healthcare system. Research-based treatments that result in shorter and/or more effective treatments may also result in savings in non-medical direct costs such as custodial care, special diets, tutors, transportation, special equipment, government and voluntary community support programmes.<sup>10,46</sup>

Simply focusing on healthcare savings may be too narrow a perspective, and some studies also look at the benefits, and/or indirect cost savings, in terms of avoidance of lost production. Using the human capital approach, which essentially values health gain in terms of the value of productivity gains, Mushkin<sup>63</sup> attempted, despite data problems, to calculate the economic benefits to the USA of all health research. In a series of calculations she (1) estimated the economic value of the total reduction in mortality and morbidity in the USA between 1930 and 1975 and (2) estimated the value of the share of these caused by biomedical research, and, after taking away the cost of the research, produced a rate of return of 47%. In the studies considered in this review, the approach has been applied to more specific bodies of research.

The details of the approach used in each study are given in the accounts of the empirical studies (see

Table 1 and Appendix 4). Methods tend to be primarily desk analysis (e.g. decision analytic models), synthesising available information, with prospective or assumed impacts from technologies (the subject of research) estimated. A number of these studies focus on the human capital approach to indicate a rate of return (or CBA) and others include both types of cost savings. The CBA approach can also go wider and include assessment of the monetary value of the economic benefits from commercial exploitation of medical research.

### **The framework developed in the ESRC project on the non-academic impact of socio-economic research**

A project for the ESRC in the UK attempted to develop practical approaches to assess the impact of ESRC-funded research, including in the health field, on non-academic audiences.<sup>45,64</sup> The approach is based on the interconnection of three major elements: the types of output that can be expected from research; the channels through which their diffusion to non-academic actors occurs; and the forms of impact. Various permutations between outputs, forms of impact and diffusion channels result in many different ways – impact mechanisms – in which economic and social research can have an impact beyond the academic world. The outputs are the generation of codified or tacit forms of knowledge and the development of skills. Articles and reports are the major form of codified outputs. Social networks constitute a key diffusion channel in the communication and application of research. Publication is also seen as a diffusion channel, as is the mobility of researchers to employment elsewhere in the economy and the provision of consultancy by academics. The third element is the impact or forms of research utilisation. These include the use of skills and tacit knowledge to solve specific technical problems and the use of codified knowledge in the form of models that can also be used to solve problems – this can be direct or indirect – and the use of research results to influence decisions either policy or technical, and this can include confirmatory and justification roles.

Two main strands in identifying the impact mechanisms emerge from the permutations between research outputs, diffusion channels and forms of impact. First, data to identify networks can be obtained from interviews whose objectives are to map the networks of researchers and users, distinguishing those that exist before the launch of the programme from those

created as a result of the programme. In their recommendations, they pointed out that although questionnaires are cheaper than interviews, they may fail to clarify the objectives of some of the questions, and their relevance, because of the lack of opportunity for real-time dialogue. Although they saw face-to-face interviews as the best solution, the cost and time-consuming nature of these meant that they recommend telephone interviews. Interviews with initial users named by researchers followed. These used the snowball approach with users asked to identify further users, thus revealing linkages within the user community.

The second strand involves tracing the post-research activity of researchers, either in outside employment or through consultancy.

Molas-Gallart and colleagues<sup>64</sup> thought it was unnecessary to have all the background analysis that is inherent in the payback framework: “We are not concerned, when assessing impact, about the way in which the research subject was defined; that is, whether it responded to the requirements of a potential ‘client’ (problem-driven), or to intellectual curiosity (research-driven) or to a mixture of both” (p. 18).

### **The framework developed in the study of the Royal Netherlands Academy of Arts and Science<sup>50</sup> to assess the societal impact of applied health research (partly informed by methods from Spaapen and Sylvain 1994<sup>53</sup>)**

A framework was developed by the Royal Netherlands Academy of Arts and Science<sup>50</sup> to assess the societal impact of applied health research. It partly drew on an earlier Dutch project<sup>53</sup> that had developed a framework to assess the societal quality, and hence the potential value to society, of research in general, but for which the two case studies undertaken to test it were of health research centres. The Royal Academy report argued that assessing the societal impacts of applied health research should be complementary to the evaluation of scientific quality. It should be included in the national quality assessment system used for research in The Netherlands and a single external review committee should evaluate both scientific quality and societal impact. Research institutes should be asked to describe in the self-assessment report the indicated output they consider relevant in realising their societal mission: “The criteria and indicators used for an evaluation will always be

dependent on the mission of a specific research group or institution” (p. 18). Therefore, four elements are: the research team’s/organisation’s mission; its performance in relation to that mission; the prospects for the future; and recommendations for adjustments.

The report claims that “Societal impact of health care can be divided into (1) relevance for healthcare providers and the process of healthcare delivery; (2) relevance for policy makers and the process of designing, implementing and monitoring policy decisions. In assessing the societal impact of research it seems prudent to discriminate between potential (*ex ante*) and realized (*ex post*) societal impact” (p. 17). “A final objective is, of course, that the health outcomes of individuals and the population would be improved. However, evaluating this would be a long-term challenge, requiring extensive specific research. In a more general societal impact assessment, therefore, indirect indicators are to be used” (p. 18).

Having set out societal impacts as the two points above, the framework provides a list of output categories and gives a qualitative description of indicators for possible applications. It suggests that the actual criteria should be elaborated in advance for specific assessment and the evaluators consider an initial self-assessment, consulting stakeholders if necessary. In the overall assessment it is **the total profile** of a research project, programme or institute **in relation to its mission** that is important. The indicators listed vary in the extent to which they are routinely available for the researchers to gather and present to the evaluation team in their self-evaluation.

The criteria (in bold) and the relevant indicators are listed here:

- **content analysis:** professional publications, treatment guidelines and protocols, policy documents, Cochrane library, textbooks, teaching materials, lay publications, ICT and software
- **citation analysis:** scientific publications [both Institute for Scientific Information (ISI) and non-ISI] in documents mentioned above, citations in professional journals, policy documents, protocols and guidelines
- **authorships:** authorships of documents mentioned above under content analysis
- **products:** healthcare technologies and services, instruments programmes, methods for assessment/implementation of care
- **funding of research:** (semi)governmental funding
- **publicity:** presentations for non-scientific audience, fact sheets, public media, Internet
- **memberships:** member of committee issuing a policy document or a treatment guideline, member of advisory committee
- **teaching:** contributions to education of healthcare professionals based on research output
- **implementation strategy:** membership of advisory committees, interactions between researchers and public administration; feedback from target groups
- **independence:** operationalisation of research questions, research methodology, analysis and publication of results.

It is clear that this framework is intended to be applied more to the work of specific teams of researchers than necessarily to the research produced in particular programmes.

The prime goal of Spaapen and Sylvain’s 1994 study<sup>53</sup> was “to develop a methodological framework through which scientific and technological research can be assessed with respect to its potential value for society” (p. i). A summary of the integrated framework is presented: (1) criteria of societal quality based on the interaction between research and societal context leading towards; (2) charting of the particularities of the communication structure, including output, media (i.e. mechanisms for transmission) and demand categories; (3) content oriented analysis of main research issues and societal demand in particular cases; (4) concrete and strategic indicators to measure different elements of (2) and (3); (5) composition of suitable juries; (6) consultation with involved actors about (1)–(5). The framework has three main assessment categories: (1) different forms of research output (written and non-written); (2) different forms of societal demand (concrete and strategic); and (3) different media through which (1) and (2) interact (e.g. journals, conferences). At one level this conceptual approach addresses the complexities of research impact; it does not, however, go as far as considering the impacts in terms of impact on policy, practice or health sector benefits.

Although the approach developed for the ESRC was unconnected to this framework, there are certain similarities between the approaches, with both focusing on types of outputs and mechanisms of diffusion or transmission.

### **Framework presented by Lavis and colleagues (Canada) (e.g. Lavis *et al.*<sup>43</sup>) to assess the impact of health research on policy (with a focus on a specific target audience)**

A major stream of work assessing the impact of health research on policy has been conducted in Canada by Lavis and colleagues and many of the approaches are drawn together in a paper by that group.<sup>43</sup> It starts by listing various assumptions. These include the view that impact measures should be based on an evolving understanding of how best to transfer and facilitate the uptake of research knowledge and should enable meaningful assessments within peer groups that fund or produce similar types of research knowledge for similar types of decision-maker. The use of research knowledge to inform decision-making, not a change in health status, constitutes the most appropriate generic measure of the impact of research that can be expressed routinely. Cultural shifts that would facilitate the ongoing use of research knowledge in decision-making, such as the creation of a research-attuned culture among decision-makers, constitute the most appropriate generic measure of the impact of future research. Impact measures can be categorised according to whether the active role in promoting research is played by researchers (producer-push measures), decision-makers (user-pull measures) or both researchers and decision-makers (exchange measures). These measures can be categorised according to whether they capture a process associated with the pursuit of research impact, an intermediate outcome, such as a change in awareness, or an outcome, such as selecting one policy option over another. Research knowledge may be used in instrumental, conceptual or symbolic ways.

According to the framework, different target audiences warrant different measures of impact. Some research is produced and transferred as a basic contribution to knowledge, in which case the target audience is researchers. The target audiences for applied research include the general public, patients, clinicians, managers, development officers in companies and public policy makers. The types of decisions that can be informed by applied research knowledge vary by target audience.

Building on the above, the framework sets out various steps that should be adopted:

- identify target group
- select appropriate measures depending upon who has been actively promoting research use

among these target audiences (producer-push/user-pull/exchange)

- select measures according to resources available (process if limited funds; intermediate outcomes if sufficient resources to conduct a survey; outcome measures if funds for case studies that can assess whether knowledge used in decision-making)
- then identify data sources or collect new data, analyse the data and identify areas for improvement and feed this back to those involved.

A small section of the many specific methods identified is listed here:

- For producer-push process, measures include the number of products published and the number targeted at specific decision-makers, for which the data sources include, respectively, CVs, impact factors and half-lives, and research organisations' websites.
- For producer-push intermediate outcomes, measures include decision-makers' awareness of research and its source, and data sources are surveys of, or structured interviews with, decision-makers.
- For producer-push outcomes, measures include decision-makers' use of research, and data sources include surveys or structured interviews with decision-makers, documentary review.

A full account of the detailed methods available for possible use is given in the paper, which provides a menu of options.

### **Organisational theory approach/framework presented by Hailey<sup>37</sup>**

Ways of assessing the impact of health technology assessment programmes were addressed as part of a wider attempt to develop self assessment tools for health technology assessment programmes to help identify opportunities for future investment.<sup>37</sup> It describes an organisational theory approach to measurement of health technology assessment programme effectiveness and examines four models: the Goal, Constituency, Open Systems and Competing Values models, and finds that all have factors that can be taken into account, but none appears to be entirely applicable to a health technology assessment programme.

The recommended approach gives considerable attention to the processes and the outputs: the "immediate impact of an HTA on the decision-making process is often difficult to determine.

Usually the HTA will be only one of a number of factors influencing a decision. The final stage in the process, consequent changes to healthcare and/or health outcomes may have a tenuous link to the HTA". It therefore concludes that, "Measures of effectiveness ought to include at least some consideration of activity, continuity, relevance, quality and impact". In relation to impact, the relevant items include: whether the report is considered; whether the recommendations are accepted and material from an HTA product is incorporated into policy or administrative documents; and whether information in an HTA is used as a reference for future activities, e.g. in subsequent development of guidelines. In relation to indirect impact, items include: influence on others in healthcare systems and broader influence on other healthcare systems.

The report highlights the need to consider how information on the effectiveness of a health technology assessment programme is to be used, and that the effectiveness of a health technology assessment programme is determined by many factors, including downstream components that are largely not under the control of the programme. An overview of the Alberta Heritage Foundations for Medical Research (AHFMR) health technology assessment programme is presented as a case example of how the health technology assessment effectiveness profile might be applied, but this does not really amount to an empirical study. It has also informed later analysis.<sup>59</sup>

### **Approach presented for assessment of health technology assessment policy impacts and cost savings in Quebec, Canada (e.g. Jacob and McGregor),<sup>40</sup> using detailed case studies and follow-up analysis**

A series of studies of the benefits from health technology assessments conducted in Quebec by the Quebec Council on Health Care Technology Assessment (CETS) adopted a conceptual framework that focuses on impacts on policies and resulting direct cost savings.<sup>39,40</sup> Although the approach was developed further in the second study, the basic approach was the same. A key step was the use of case studies based on documentary analysis and interviews to identify, for each subject addressed, "the dynamic of the interactions that took place between the HTA in question and the positions taken by deciders". These started with careful scrutiny of documents including administrative files and public pronouncements of the Ministry and committees and other public documents. Traces of influence in these

documents were called "critical incidents". To be considered a critical incident, it was essential that evidence of influence of a health technology assessment found in these documents should refer explicitly to the health technology assessment in question and be used to support actions that were consistent with its content. Critical incidents were also identified in advisory letters or normative guidelines issued by the professional organisation charged with this responsibility, the College of Physicians of Quebec.

To complete and validate the evidence derived from these sources, contact was made with the key actors. Critical incidents were classified into seven levels of importance according to the type of policy involved, such as general statements of ministerial policy, planning guidelines and ministry decisions on allocation of resources. Taking into consideration the level and number of critical incidents, an overall estimate of impact on policy was awarded to each HTA on a four-point scale from 0 (no impact) to +++ (major impact). The weight awarded to critical incidents was adjusted according to the nature of the decision, for example, for decisions relating to health service coverage, for impact to be rated "substantial" (++) at least one critical incident had to be found at level 4 or above. Impact on policy was considered zero if only level 1 incidents were found, and an extra + was awarded when the case study revealed a clearly dominant influence of the health technology assessment over other sources of influence. In addition to searching for critical incidents, the impact of the health technology assessments on the decisions of healthcare professionals and hospitals was also documented through utilisation data and by questionnaires.

The impact on costs was measured by estimating the difference between how much would have been spent on a technology in any year if no health technology assessment had been written compared with what was actually spent, the former estimate being based on projections of trends in the years preceding the health technology assessment. Next the proportion of the difference between projected and actual costs that should be attributed to the health technology assessment was estimated, using upper and lower limits.

### **The knowledge utilisation model developed in Alberta, Canada (health technology assessment research of AHFMR)**

An assessment was made of the 2000–1 products from the health technology assessment

programme of the AHFMR.<sup>16</sup> It examined context (environmental enablers and barriers that influenced impact), implementation (critical processes and activities in producing health technology assessment products that influenced impact), and outcome (extent to which reports influenced healthcare policy and decision-makers).

A more extensive impact assessment of the 2001–2 products was undertaken with a more sophisticated model of knowledge utilisation adopted.<sup>17</sup> It attempted to go beyond ‘instrumental utilisation’. In addition to the impact considered in the 2000–1 report, interviews were conducted with the hope of identifying impact in terms of enhanced knowledge, new research questions and ideas and improvements in the health of Albertans and the quality of health services in Alberta. There was also an intention to determine the degree to which the questions asked by requesters of health technology assessment products were answered and learn if there were unexpected outcomes that came about as a direct result of the involvement of the health technology assessment unit with a requester’s organisation.

### **Impact model developed by the USA Agency for Healthcare and Research Quality (AHRQ)**

This model was developed and described by Stryer and colleagues<sup>54</sup> following data collection via a survey for an assessment by the AHRQ of the impact of one of its programmes. The model was later seen as being more widely applicable for AHRQ.<sup>30</sup> The model depicts potential levels of impact. It is presented as a triangle with research findings being at the base and constituting the widest level (level 1); next, level 2 is impact on policies; level 3 is impact on clinical practice; and level 4 is impact on healthcare outcomes. Level 1 impacts represent the foundation of studies that identify problems, generate hypotheses, establish the effectiveness of interventions, and develop new tools to explore these problems. These studies include findings which conflict with current clinical consensus and practice. Level 2 impacts are those in which a policy or programme is created as a result of the research. Potential conduits of level 2 impacts include various change agents: health plans, professional organisations, legislative bodies, regulators, accrediting bodies, the media, industry and patients. Level 3 impacts are those in which there is a documented change in what clinicians or patients do. Level 4 impacts are changes in health outcomes, including clinical, economic, quality of life and satisfaction.

## **An introduction to the empirical literature on assessment of health research programmes**

Examining the empirical literature describing the application of the models identified, plus other studies, provides details of how assessments work in practice. *Table 1* presented an outline of each of the 46 studies included in our body of evidence. Of these, 41 were classified as empirical studies.

Appendix 4 presents a more detailed account of the empirical studies linked to each of the conceptual approaches plus a selection of the other studies that are most relevant for the assessment of the impact of the NHS HTA Programme.

However, the literature continues to develop, and although we are not able to cover the most recent contributions in detail, some key recent contributions will be highlighted in the section ‘Summary and conclusions’ (p. 43). These include: an evaluation of the health gain and monetary value of the clinical trial from one of the National Institutes of Health;<sup>65</sup> an evaluation of the gains in knowledge, health and wealth from research funded by the Australian Medical Research Council at various periods;<sup>66</sup> an attempted application of the payback framework to assess the impact of population health surveys in Australia;<sup>67</sup> an assessment of the use made by the legislature of the research from the California Health Benefits Review Program;<sup>68</sup> the development of a new model for assessing impact and its application to some research conducted at the London School of Hygiene and Tropical Medicine;<sup>69,70</sup> and the application of the payback framework in an assessment of the impacts from the HTA Programme of ZonMw<sup>71</sup> and publicly funded health research in Hong Kong.<sup>72</sup>

## **Summary of key findings from studies that assess the impact of entire multi-project programmes**

Although the findings from the included studies are not the main focus of this review, and cannot be analysed in any statistical way, it might be useful to attempt to draw together those findings from the studies that most resemble the impact assessment being conducted of the UK’s health technology programme. *Table 2* lists the studies that attempted to analyse all the projects within a multi-project programme. As with other aspects of



**TABLE 2** Summary of key findings from studies that assess the impact of entire multi-project programmes<sup>a</sup>

Authors	Date	Main techniques	No. of projects/ response rate	Average no. of articles	Impact on policy (%)	Impact on practice (%)
<b>Health technology assessment programmes</b>						
AHFMR <sup>17</sup>	2003	Ints	10: most		Most	
Hailey <i>et al.</i> <sup>33</sup>	1990	Desk	20		70	
Hailey <i>et al.</i> <sup>36</sup>	2000	Ints	20: all		70	
Jacob and Battista <sup>39</sup>	1993	CS + desk	10: all		80	
Jacob and McGregor <sup>40</sup>	1997	CS + desk	21: all		86	
McGregor and Brophy <sup>44</sup>	2005	Not given	16: all		100	
Shani <i>et al.</i> <sup>52</sup>	2000	Commentary	84: all		88–100	
<b>Other health research programmes</b>						
AHFMR <sup>17</sup>	2003	Q + Ints	100: 50%	1.1	49	39
Buxton <i>et al.</i> <sup>24</sup>	1999	Q + desk + CS	167: 70%	1.1	41	43
Ferguson <i>et al.</i> <sup>32</sup>	2000	Q + desk	119: 75%	1.9	16	25
Hanney <i>et al.</i> <sup>11</sup>	2003	Q + desk	36: 83%	2.0	30	27
Molas-Gallart <i>et al.</i> <sup>45</sup>	2000	Ints	15: 93%		50	
NHSE Trent <sup>48</sup>	1997	Q	37: 80%		<10	<10
Shah and Ward <sup>51</sup>	2001	Q + desk	55: 69%	1.6	53	69
Stryer <i>et al.</i> <sup>54</sup>	2000	Q + desk	115: 64%		Limited	Limited
Wisely <sup>56</sup>	2001	Q + desk	63: 63%	2.3	35	27
Wisely <sup>57</sup>	2001	Q + desk	39: 67%	2.1	31	23
CS, case studies; Ints, interviews; Q, questionnaires.						
<sup>a</sup> Health technology assessment programmes are broadly health technology assessment; other health research programmes are broader but include some health technology assessment.						

this review, it is difficult to draw precise inclusion criteria. The studies included in *Table 2* considered at least some elements of downstream impact, but there is great variation in subject matter, including mostly responsive mode laboratory-based work in the NHSE Trent study, and variations in time elapsed since projects were completed – both within and between programmes. Furthermore, the table inevitably greatly simplifies what is meant by terms such as impact on policy. Many of the questionnaire studies rely primarily on researchers' self-reporting, and, at least in those involving the payback framework, a broad interpretation of impact on policy was encouraged.

It is also difficult to classify these studies in a way that will provide appropriate comparators for the NHS HTA Programme. The included studies assessed the impact of a wide range of

programmes and any attempt to classify them into two groups is an oversimplification of what is in reality a spectrum. Nevertheless, *Table 2* attempts to distinguish between health technology assessment programmes and other health research programmes, but some of these other programmes contain health technology assessment projects. Furthermore, most of the health technology assessment programmes listed here are ones that fit into the category, as discussed in Chapter 1, of being essentially reviews linked to the needs of specific organisations.

The diversity of the underlying programmes covered by the studies included in *Table 2* is matched by the diversity of approaches used in those studies to assess impact. This again limits the extent to which comparisons can be made. In some instances, such as in Hailey and colleagues,<sup>33</sup> the analysis focused mainly on the technologies

assessed than on the specific reports produced, but in most other studies the focus was on the pieces of research (or projects) that constituted the programme. Over half of the studies included in the other health research programme category are based on the Buxton/Hanney payback framework.

Reflecting the diversity of both the underlying research programmes and the approaches used in the impact studies, two of the columns, 'Average number of articles' and 'Impact on practice', are blank for the health technology assessment programmes. Most of these health technology assessment programmes are of the type where reports are of key importance and the impact studies focused mainly, though not exclusively, on the impact on policy made by the reports.

## Discussion

### Limitations including identifying the included literature

The identification of the relevant literature was a complex task given the undeveloped and idiosyncratic nature of this field. The intention was to concentrate on the body of research most relevant for the task of assessing the impact from the UK's NHS HTA Programme.

It is relevant to consider how far the type of research being included influenced the techniques and findings that emerged as important in the review above. Hanney and colleagues<sup>11</sup> identified possible differences between, first, studies that start with a body of research and examine its impact and, second, those that consider developments in the health sector, especially policy decisions, and analyse how far research, from whatever source, influenced these developments. The exclusion of the latter studies from our final selection not only means that some prominent studies in the overall field of the impact of research, for example that by Lavis and colleagues,<sup>73</sup> are not included, but also possibly helps to account for the generally more positive picture that emerges about the degree of impact that is made by research. This is because, as a generalisation that is discussed below, studies that start with specific research have the advantage of a more tightly defined focus that can help in identifying that the research has made some contribution.<sup>11</sup>

There are also issues around how far it is appropriate to include studies where the health technology assessment study itself was more like

an appraisal from an official body than an assessment undertaken as a research study. In some cases, appraisals were excluded from this review even though some appeared in the review of the impact of health technology assessments conducted by Gerhardus and Dintisios.<sup>9</sup>

In composing the body of evidence, we excluded papers that provided just a commentary, as opposed to a model and/or an assessment application. This might mean that some of the contributions most critical of the whole approach of assessing research impacts have been excluded, but at least some of the comments are brought into the analysis below of strengths and weaknesses of various approaches.

### The nature of the included body of evidence

Comparatively few studies were identified that had attempted to develop a general conceptual or methodological framework to assess the benefits. We were therefore pulling together a very disparate range of studies, most of which were undertaken for the purpose of assessing the impact of a specific body of health research and only a few of which had more methodological purposes.

In this field, there is no single standard approach and many of the studies have been one-offs. Many of those one-offs were applications for which there was no real attempt to develop a model. Various other studies started with the task of assessing the impact from a specific programme of research and developed a conceptual framework either during the study or after the data had been collected. Examples of the first include the health technology assessment policy impacts and cost savings in Quebec. Examples of the second include the study by Stryer and colleagues<sup>54</sup> for the AHRQ in the USA.

It is possible, however, to identify a few frameworks that were specifically devised for assessing the impact from either health research or social/applied research in general, and have been tested by being applied to a body of research. These include the payback framework, assessment of the non-academic impact of social research for the UK's ESRC and the assessment of societal research in The Netherlands. In yet other cases, the conceptual development was not accompanied by an application, for example the Canadian comprehensive framework for measuring the impact of health research. Finally, at least one conceptual framework,

CBA, is used much more generally in other fields but has been adapted for assessing the impact from research.

The question of how far a conceptual framework is necessary, or desirable, can be informed by the listing of, first, conceptual approaches, then empirical studies and then a discussion of the impact assessed. Possibly all that is needed is a set of clear objectives such as the assessment of impact across several categories and explicit techniques or methods to gather the data on impacts. There are some important studies<sup>32,51</sup> that fall into this category. As noted above, however, Ferguson and colleagues<sup>32</sup> claim that their study was informed by the existing conceptual approaches such as the payback framework.

In terms of empirical studies, lessons can be learnt from the broad body of evidence included. It covers some studies where the 'programme' of research is restricted to a few studies on the same topic, or one major stream of research in a particular field or the work of research centres. The topic of the research programme also varies widely, with perhaps 14, or about one-third, of the empirical studies being related to health technology assessments. Even here, however, the diversity was great because only a few of these were based on whole health technology assessment programmes; others were based on a selected number of health technology assessment projects or on the work of a body that was at the borderline between having responsibility for conducting health technology assessment research assessment studies and being the official body that conducted formal appraisals. Some studies concentrated on impact, however defined, others had a broader evaluative focus. In terms of using the review to draw out findings and lessons for the completion of our own impact assessment, probably the most useful studies are those that focus on the assessment of a whole multi-project programme. These are listed in *Table 2*.

One issue that becomes clear from *Table 1* is frequent conflicts of interest involved in impact assessment, particularly because most evaluations are sponsored, or funded, by the body that originally funded the programme of research.

Different studies focus on different types of impact from research. For some of the studies, the types of impact being analysed form an integral part of the conceptual or methodological approach. The various payback studies, for example, use the

multi-dimensional categorisation of benefits described above. However, some other studies also use a similarly broad range of impacts but do not specifically build them into a framework. Where a wide range of impacts is included, as with the payback framework, they include some items such as knowledge production and research training that form more traditional elements of research evaluation with its concentration on scientific quality. Studies that only focus on such items have generally been excluded from our study as they do not really count as impacts.

Some of the included studies use particular combinations of impacts, for example on policy/decision-makers and cost savings. Others focus on one specific type of impact, particularly impact on policy or impact in terms of a monetary value. In addition to the focus on impacts, many studies also analyse various process issues.

All eight payback studies naturally use the multi-dimensional categorisation that is an integral part of the HERG payback framework. In only a few of these studies, however, was there much of an attempt to identify, and provide evidence of, the broader economic benefits that come from a healthy workforce or from commercial development of research-based products.<sup>74,75</sup> Even in these studies the evidence about broader economic benefits was limited. A wide range of studies took a broad approach to impacts, other than economic ones.<sup>15,17,21,22,28,30,32-35,41,42,45,48,51,53,54</sup> Similarly, the new study by Kingwell and colleagues<sup>66</sup> had a broad focus.

In the Quebec studies, the impacts on policy and on cost savings were the main features.<sup>39,40</sup> In other studies, the main, or sometimes sole, impact being considered, apart from publications and training, is on further research or on policy (broadly interpreted<sup>11</sup>) or on CBA/monetary benefits. For further research the studies include that by NHSE South West.<sup>47</sup> For policy there are various studies<sup>16,17,19,20,26,36,38,58</sup> and the new studies by Oliver and Singer<sup>68</sup> and by Anderson.<sup>67</sup> For CBA there are four studies.<sup>29,46,49,55</sup>

Much of the discussion in this sub-section has been organised around the HERG payback categories. Some of the studies use rather different terms and a potentially important question then becomes whether the multi-dimensional categorisation of the payback approach is missing out any important impacts or just presenting them in a different, more structured way.

## Findings from the review

Considerable impact has been reported from studies that looked at whole programmes of research (*Table 2*), much more than is usually thought. Why might this review have identified a greater level of impact than is often reported? What reasons have been identified in the studies for impact being made by these programmes?

Gerhardus and Dintsios<sup>9</sup> also reported considerable levels of impact of health technology assessment, especially on policy making, but they pointed out that most of the studies that they included were conducted for, and sometimes by, the health technology assessment agency itself. Possible conflicts of interest were also identified in many other studies. Such conflicts are more likely in studies of impact that start with a body of research, compared to those that start with a policy or practice development in the health service (see below). Similarly, different types of evidence are sought in studies which start with the research. Many of these studies (although not some of those from AHFMR or Quebec or the early ones from Hailey and colleagues), rely heavily on self-reports from the researchers. Starting with researchers and tracing forwards does provide a tighter focus that helps with identifying any contribution that the research might have made.<sup>11</sup>

The fact that most of the programmes of research included in *Table 2* are health technology assessments or other applied research could also potentially help explain the higher levels of impact. Although the degree of impact of health technology assessment programmes has been described as variable in some previous reviews,<sup>11,76,77</sup> the level of impact for the health technology assessment studies included in *Table 2* mirrors Gerhardus and Dintsios<sup>9</sup> in being consistently high. Hanney and colleagues<sup>11</sup> discussed the importance of the receptor (or customer) role in enhancing the likelihood of a programme of research making an impact. Such factors were identified as playing a part in several studies, including those from AHFMR, Quebec and Israel.

## Strengths and weaknesses of the various approaches

The strengths and weaknesses of different approaches can vary depending on the reasons for assessing impacts. The sources drawn upon for this analysis go wider than the specific papers included in even the wider body of evidence. This is because the analysis of strengths and weaknesses often comes in commentaries, sometimes from

outside the health field. We consider questions that will face any particular assessment before looking at specific approaches described above that were applied in more than one study.

The key questions facing any impact assessment include:

- Should the level of analysis start with that of individual projects?
- Should the direction of analysis be tracing forwards from research?
- Should there be a conceptual framework to organise the assessments?
- Should a multi-dimensional categorisation of impacts be adopted?

### Level of analysis

If the objective is to assess the impact of a multi-project programme, such as the NHS HTA programme, a strength of focusing the assessment on individual projects would be that it provides a comprehensive way of gathering data at the level of the basic units that make up the programme. Where the programme consists of several types of projects or modes of funding it can be valuable, as with the study by Wooding and colleagues<sup>75</sup> to consider the impact made by projects in different funding streams.

There are, however, practical and theoretical objections to approaches that start with the projects. Gathering data about each project can be resource intensive and some researchers are unhappy about their work appearing to be assessed on criteria over which they might have limited control.

Starting with projects implies a linearity that may obscure the complex reality of how policy and practice changes occur. This criticism, sometimes called the 'project fallacy', is made widely in debates about assessing research impacts.<sup>78-80</sup>

Studies that focus on particular projects may give insufficient attention to general outcomes. Rather than looking at the impacts from individual projects, it can be asked how far the research programme has helped generate a cultural shift towards the adoption of evidence amongst policymakers and practitioners in the healthcare system. Lavis and colleagues<sup>43</sup> claim that cultural shifts constitute the most appropriate generic measure of the future impact of research.

### Tracing forwards from research

In assessing health research impacts, two broad approaches have been identified.<sup>11,81</sup> These involve

starting with research and tracing forwards to identify its impact or starting with impacts and trying to trace back to the research behind them. The strengths of tracing forwards from specific research are that it provides for more focused studies and this might be particularly relevant where the assessment is focused on the impact of a specific research programme. As noted above, there are indications that such an approach can make some headway in identifying that many projects can make at least some contribution to policy making. There is a clear link between this approach and focusing the assessment at the level of projects.

At the same time, it might be a weakness of the approach that there is a tendency to exaggerate the impact of the specific research under consideration. This is because such studies can appear to ignore the difficulty that any impact a research study makes is usually of a contributive nature and it is difficult to quantify the impact of one project from that of others with which the findings get mixed. Overlapping with these weaknesses are ones that flow from the techniques most often used in forward tracing studies, that is, questionnaires. These weaknesses include the heavy reliance on self-report data from lead investigators who may indicate impact where they hope the findings will have impact, rather than when actual impact has been seen, ambiguity in response data in relation to how policy decisions are defined in the questionnaire and the cross-sectional nature of most applications of such questionnaires, which can cause problems for the most recently completed projects. In general, conflicts of interest, as noted above, are more likely to occur in studies that trace forwards.

It is sometimes thought to be more realistic to start with the perspective of the policy or practice changes and consider the full range of factors that might have been responsible for such changes.

Major examples of studies in this mode that adopted a bibliometric approach (in our key references, but excluded from the body of literature due to not relating to a specific programme of research) include the classic study by Comroe and Dripps<sup>82</sup> and the work inspired by it by Grant and colleagues.<sup>83</sup> It is difficult to see, however, how such an approach could be applied across the board to assess a diverse programme such as the NHS HTA Programme.

#### **Use of a conceptual framework**

A framework can help organise the conduct and presentation of studies by identifying factors

linked to impact and by facilitating comparisons between case studies. A conceptual framework can structure the data collected whether in documentary analysis, surveys, interviews or case studies. A conceptual framework based on research utilisation should focus on specific issues, for example, dissemination mechanisms.

A conceptual framework may, however, introduce unnecessary rigidity into the data collection. Extra costs may be involved in collecting additional data but these may be offset by avoiding the gathering of irrelevant data.

#### **Adopting a multi-dimensional categorisation of benefits**

In terms of the categories of benefits on which to focus, the above discussion clearly reveals that there are many options. Some studies adopt a broad approach to impact assessment, as seen in studies such as those by AHFMR,<sup>17</sup> Kingwell and colleagues,<sup>66</sup> Shah and Ward<sup>51</sup> and Stryer and colleagues<sup>54</sup> and also those studies explicitly using the Buxton and Hanney payback approach. This approach is compatible with the stated aim of many health research funders and researchers. For example, the recent document outlining the response to the consultation on the new National Health Research Strategy in the UK states that in terms of judging the impact of the research strategy, "With regard to patients and the public, improved care outcomes were seen as the crucial measure and respondents thought it was important to try to assess these impacts".<sup>84</sup>

This approach has been questioned by experts such as Lavis and colleagues,<sup>43</sup> Hailey<sup>59</sup> and the Royal Netherlands Academy of Arts and Sciences,<sup>50</sup> who argue that it is impractical, apart from in detailed case studies, to go beyond examining impact on policy making. At the empirical level, some studies, such as that by AHFMR,<sup>17</sup> attempted to look for wider examples of utilisation, but in practice found resource constraints limited their ability to identify users beyond the immediate requesters of the health technology assessments.

Some studies, including that by Stryer and colleagues,<sup>54</sup> that adopt the comprehensive approach find it does not capture the general contribution of the programme's work: specifically, the AHRQ's programme on Outcomes and Effectiveness Research may have played a major role in the expansion of interest and capacity in outcomes research that was not captured. Rather than attempting to assess impact on policy,

practice and health gain across a whole programme, it may be more appropriate to focus on the 'process' indicators that show the degree of communication and collaboration between researchers and potential users of research.<sup>43</sup> This would link with attempts to show how the research funding body had helped to generate a change in culture.

We next consider these points in relation to the three approaches that we identified as having been applied on more than one occasion: CBA, the Quebec approach and payback.

CBA, or applying a monetary value to the benefits from research, was applied in a series of studies. Possible strengths include the clear and widely accepted conceptual framework that can be applied independently of the body that funded the research. The major strength of this approach is that if it can be operationalised, it does provide a single monetary figure for impacts that can then be compared with the benefits from other research and indeed other forms of public expenditure. Operationalising the approach, however, is not easy. Such studies face all the problems associated with having to show that the specific research has led to the health gains claimed and then an acceptable approach has to be found to valuing the health gain. There are now substantial criticisms of the human capital approach which is often used.<sup>10</sup> Furthermore, it could be suggested that none of the monetary valuation studies discussed earlier in this chapter focused on a whole multi-project programme of research but instead on selected bodies of research. This criticism is now less valid following the study by Johnston and colleagues.<sup>65</sup> The implications of this new study, and how it could help inform future activity in the UK, are discussed later in this report.

The conceptual approach developed in Quebec for assessing the impact of the HTA programme in terms of impact on policies, and subsequent cost savings, informed studies that produced evidence of substantial impacts. The strengths of this approach include its tight focus on the impact made by a series of specific studies within a programme, and doing so without having to rely just on data from researchers. It is likely that the impacts were enhanced by the particular organisational arrangements for health technology assessment in Quebec which the studies also explored. A possible weakness is that the figures for cost savings were impressive but questions are raised about how far the savings are actually

realised. These studies do seem to have been broadly successful in addressing the questions they were asked and in the context in which they were undertaken. It is important to note, however, that the studies being examined took the form of reviews specifically aimed at informing policy in a relatively small healthcare system that was receptive to such an approach. The detailed approach might be less feasible in other circumstances.

The most commonly used framework, the payback approach, has various strengths. The frequent adoption of a combination of methods (questionnaire/interview/desk analysis) across some key payback dimensions, and using the payback model as a research sequence, present as a useful and feasible approach to consider the assessment of impact of the NHS HTA Programme. The general advantages of having a conceptual framework were discussed above and, in particular, the payback framework builds on understanding and modelling about how research is utilised in policy making and practice – especially the major study on the organisation of needs-led health research in the UK.<sup>62</sup> The payback framework has been applied in a variety of settings and is very amenable to being applied through the use of questionnaires where there are a large number of projects. As noted above, the combination of a multi-dimensional categorisation of benefits and the various stages of the payback model have been useful in organising the conduct and presentation of studies in a way that facilitates comparisons and helping to identify factors that might be linked to having impact – both in questions asked in surveys and in interview schedules.

Weaknesses include the difficulty in identifying all the payback dimensions. Some of the criticisms noted above (to do with the weaknesses of focusing at the project level, tracing forwards and attempting to identify impact on practice and even health gain) were specifically identified in relation to the payback framework.<sup>43,80</sup> Although Molas-Gallart and colleagues<sup>45</sup> advocated a conceptual framework to help organise the assessment of research impact, they suggested that the payback framework focused on too many issues around the generation of research projects that were not necessary in such an assessment.

Some of those who identify the weaknesses of the payback framework acknowledge that it has been applied in a series of studies. The specification for the current project set out that the payback framework should be used in the assessment of the

impact of the UK NHS HTA Programme. The analysis reported here of strengths and weaknesses of various approaches was conducted concurrently with the actual impact assessment and therefore fed some ideas into the study in terms of how the approach could be improved.

### Lessons from the review for our impact assessment and issues for further research

Various lessons are considered in terms of issues of potential relevance during the conduct of the assessment of the impact of the NHS HTA Programme. A variety of techniques have been used to gather data, but the only studies that did not use questionnaires as a key element were those where the number of projects examined was fewer than 30 (in the case of Shani and colleagues,<sup>52</sup> the study was more of a commentary on the overall work of the Medical Technology Administration than an assessment of the impact of the projects). The findings of these studies were often thought to be more robust and informative where other techniques were also adopted. Molas-Gallart and colleagues<sup>45</sup> argued that telephone interviews would be preferable to questionnaires. The lesson from this is that we should assess the extent to which questionnaires can be relied upon to produce useful and accurate data.

*Table 2* indicates that the response rate for surveys was usually about two-thirds, rather less than that obtained using other methods. This raises issues about whether the responders are representative of the full population. Wisely<sup>57</sup> made an important point in her study by showing that the response rate was higher from those projects where the original proposal had received the highest ratings in the proposal review process. The lesson from this is that it is worth attempting to use any information available about the characteristics of responders and non-responders.

Despite the interest in adopting a broad approach to impact assessment, in practice it is difficult to do. Therefore, it is useful to be flexible in gathering data and attempt to use any sources that might become available. In general, the evidence from empirical studies was less strong in relation to impacts on practice and on possible health outcomes than it was in relation to impact on policy. Both the payback and the Quebec approaches recognise the relevance of trying to assess the counter-factual, that is, what would have happened had the particular research in question not been conducted. The third lesson has to do with the value of using triangulation wherever

possible, for example, using analysis of documents in addition to claims made in questionnaires and interviews.

An issue that emerges from the study by Dixon and colleagues<sup>28</sup> is how far it is important to have locally conducted health technology assessments. This was claimed to be a significant factor in several studies at the level of the nation (Israel<sup>52</sup>) sub-national units<sup>40</sup> and even local hospital units.<sup>44</sup> This again raises the desirability of considering the counter-factual and attempting to assess how far it was the specific HTA under consideration that made a difference to policy and practice.

Some studies, including that by Stryer and colleagues,<sup>54</sup> were concerned that the framework adopted does not capture the general contribution of the programme's work. The lesson for the study of the impact from the NHS HTA Programme has to do with trying to capture such views about general impact. Some general interviews were conducted in the study's pilot. The current study should include analysis of the context such as the development of interface mechanisms and receptor bodies.<sup>11</sup>

If the assessment is to provide useful information to the HTA Programme on how to achieve maximum impact, then progress will have to be made in identifying the various factors linked to achieving payback. Various suggestions have already been made. Some are the contextual issues such as the organisational arrangements in which the research programme is situated. Others are more to do with the qualities of the individual research and activities of the specific researchers. With its conceptual framework and multi-method approach, the payback framework can and will be used to address these issues.

A major final lesson is that it is appropriate to spend time on methodological considerations.

Several studies, including those by Ferguson and colleagues<sup>32</sup> and NHSE South West,<sup>47</sup> conducted bibliometric analysis and examined the journal impact factors of the journal outputs from the projects in the programme under consideration. Other studies, including those by Buxton and Schneider,<sup>25</sup> Shah and Ward<sup>51</sup> and Wooding and colleagues,<sup>75</sup> attempted to make some correlations between journal outputs and other forms of impact. This often proved difficult to achieve, but if such an approach could be developed, and if, contrary to the initial evidence, it did show

positive correlations, then this would suggest outputs could potentially be used as proxy indicators for possible subsequent outcomes. This would be most helpful as attention turns to the development of approaches to regular monitoring of impact. It partly links with the approach of the Royal Netherlands Academy of Arts and Sciences<sup>50</sup> in recommending 'indirect indicators', including presentations. Others, however, remain to be convinced about the value of statistics on conference presentations<sup>32</sup> or aggregate data on the percentage of projects that have established stable links with users.<sup>45</sup>

Other potential indicators of proxies for outcomes could be identified from the analysis above of possible reasons for the higher impact shown by some programmes. Such issues could be addressed in the analysis of the data from the study of the NHS HTA Programme, along with consideration of the appropriate timings for impact assessment on a regular basis.

Several studies, including that by Croxson and colleagues,<sup>27</sup> consider aspects of conducting regular monitoring of the impact made by a programme of research. A key issue is how far the data collection can be made less resource intensive and yet avoid the potentially distorting effects of crude indicator systems.

In terms of identifying potential comparators for the NHS HTA Programme, *Table 2* highlights the diversity among impact studies and hence the difficulty of neatly classifying our impact assessment into either of the two groups used in *Table 2*. Perhaps the NICE TARs from the NHS HTA programme are broadly comparable with the health technology assessment group because they are usually assessments explicitly produced for decision making bodies. In many ways the primary and secondary studies from the NHS HTA programme are more like many of the other health research programmes listed in *Table 2*, but even here there will be differences because not all the projects in all the other health research programmes are directly needs-led.

The diversity amongst the studies is also relevant for the possible research recommendations. There is a growing body of research in this field and, as we have seen, some important studies have been reported since the formal cut-off point for this review. This collection of studies is very different from that in most medical fields because usually each impact study is a unique analysis of a programme or body of research and it would be

unusual for another impact study to be undertaken on the same body of research. This means that the standard systematic review techniques are not appropriate. There might, however, be considerable value in a research project that collated impact studies in an ongoing manner and analysed them in a consistent fashion, even though the studies themselves would probably continue to be diverse. One of the issues that such an ongoing impact-research collation and analysis programme could consider in more detail than has been possible here is the question of the time elapsed before impact occurs, and building on that how it might be possible to assess impact over different timescales.

There is an overlapping debate about the desirability and feasibility of attempting to assess the impact of research on health gain and economic benefit. Broadly, there is seen to be a spectrum in which impact is more easily identified for knowledge production activities than for impact on policy and even more so than for impact on practice or outcomes such as health gain or economic benefit. The debate about the desirability of assessing the later parts of the spectrum is complex. Some argue that research funding bodies and systems should concentrate on those issues for which they have most direct responsibility, that is, the production of knowledge. Others argue that a health research system should be concerned with what the broad impact of its research has been and how it could be enhanced. From the perspective of informing this debate, it could be important to make progress in showing the feasibility of assessing health gain and economic benefit. This presents itself as a major challenge for future research.

One possible way of advancing such a stream of work would be to develop case study approaches to help assess the impact on behaviour and the gains in the health sector and the broader economic benefits. Johnston and colleagues<sup>65</sup> used secondary sources about the uptake and benefits of certain interventions to assess the economic value of key clinical trials conducted by the National Institute of Neurological Disorders and Stroke (NINDS). In the UK context, it would be necessary to undertake primary research using the type of payback approach described by Buxton and colleagues<sup>10</sup> to establish that the particular research was responsible for the introduction of the specified intervention that leads to the health and economic gains. Questions about the counterfactual would also need to be addressed. Such detailed case studies might explore the relative



contributions of related projects on particular topics. While the study by Johnston and colleagues<sup>65</sup> examined the impact from a set of clinical trials from the NINDS, it is possible that for programmes such as the NHS HTA Programme, any attempt to use such intensive case studies would be on a more selective basis.

## Summary and conclusions

The systematic review of the literature on impact assessment identified a relevant and interesting literature of around 200 papers. The review focused on 46 of these papers to identify useful approaches. Although a small number of these references (around 10%) were solely conceptual papers, introducing frameworks for impact assessment, the majority offered useful insights to the empirical application of the methods suggested.

The frameworks identified tended to be focused on specific impact assessment initiatives for designated research programmes or groups. However, some were more generic, applied across a range of programmes (e.g. the payback approach), or were able to be generalised to other settings. Almost all frameworks involved one or more of the core research activities of desk analysis, questionnaires, interviews and case studies.

Limitations in the conduct of systematic reviews on topics such as this include deciding on which papers meet the inclusion criteria. Difficulties also arise in interpreting the findings from studies due to their context specificity and to differences between the research programmes assessed (e.g. their scope, the period of research considered, the different methods employed and ambiguities over what is meant by impact on policy). Also of note are the potential conflicts of interest in the studies reported (often undertaken by the funding body) and possible selection biases in the studies that have been subject to impact assessment.

The studies identified and discussed were largely successful in presenting some quantifiable measures of impact from research funding. Some assessed from a multi-dimensional perspective whereas others targeted a single aspect (e.g. knowledge production or policy). Impact is more easily identified for knowledge production activities than for impact on policy and even more so than for impact on practice or outcomes such as health gain. In these latter domains, impact assessment often involves judgements on the use

made of research findings, which may involve assessing the contribution of the research to a broader body of evidence. It may be necessary to seek advice on such judgements from a broader sample than that often used in impact assessment projects (e.g. may need to contact users, policy makers, decision-makers). Studies that have solely (or primarily) focused on the impact of research on health policy (e.g. that by Hailey and colleagues<sup>36</sup>), using more detailed methods to seek opinion on policy impact, and relying on judgements surrounding level of impact have been more successful in addressing this aspect of impact from research activity.

Whilst the review presents a broad literature on approaches for impact assessment, the literature in this area continues to develop, with recent studies published by Johnston and colleagues<sup>65</sup> on the programme of clinical trials from one of the US National Institutes of Health and Kingwell and colleagues<sup>66</sup> reporting on the assessment of impact from the Australian NHMRC-funded research.

The review presented in this report adds to the general literature on impact assessment, and several points stand out: first, as noted above, the enormous difficulties in assessing the later categories of impact; second, and despite these difficulties, higher levels of impact on policy emerge compared with what has often been claimed, this may be due to the studies examined having taken the research conducted as their starting point; and third, and of particular interest in the current report, is the evidence around the use of the payback approach presented by Buxton and Hanney.<sup>23</sup> This was identified *a priori* as the framework to guide the impact assessment process undertaken in this project for the NHS HTA Programme.

From the international literature reviewed, the payback approach has been the most widely used approach to inform the assessment of impact, with 11 attributed empirical applications reported to date, and other impact assessment studies clearly drawing from the payback literature (e.g. that by Ferguson and colleagues<sup>32</sup>). The adherence, across the reported studies, of the broad conceptual payback approach has varied, providing evidence of flexibility within the approach, but also showing that in some instances the use of such a conceptual approach may not be necessary (or desirable). However, the literature does provide evidence of the acceptability of the payback approach to the research community at large, and within a policy orientated environment. It also

presents evidence to indicate that it is a feasible approach in a number of different settings. The primary tools of the payback approach can be adapted to specific assessment questions, and whilst it is clear that not all of the categories of payback defined by the approach are easily accessible in the empirical assessments undertaken to date (i.e. health sector benefits, broader economic benefits), such categories of payback are undoubtedly desirable and most commentators would agree that they should remain a target for any future assessment initiative.

In terms of informing further work, the review provided various lessons for our specific study of the impact from the NHS HTA Programme. One of the key lessons related to the complexity of

finding appropriate comparators for the NHS HTA Programme. Possibly the NICE TARs from the NHS HTA Programme are broadly comparable to the group of health technology assessment studies reported in *Table 2*. In many ways, the primary studies from the NHS HTA Programme are more like many of the 'Other research programmes' listed in *Table 2*.

The very diversity in the existing and emerging studies in this field lead to the suggestion that a research project to collate the existing and emerging studies and analyse them in various ways would be of benefit. Finally, the feasibility of assessing the health gain and economic benefits should be considered through undertaking detailed case studies.

## Chapter 3

# Evaluative framework for this study

The empirical analysis in this study used the Buxton/Hanney descriptive categorisation of payback and the payback analytical framework. This decision was made in advance of the review of the literature presented here. It reflected a recommendation in the project brief which in turn had been informed by the preparatory study led by Ruairidh Milne. Nevertheless, the literature review undertaken here, and presented in Chapter 2, informed some of the details of the conduct of the study and the analysis of findings, and confirmed the choice of approach.

### Previous use of the payback framework

The payback framework consists of a categorisation of benefits and the model (sometimes referred to as a logic model, as in Hanney and colleagues<sup>85</sup>). It was originally developed to help assess the impact or 'payback' from health services research funded by the Department of Health. The programmes of research in question included, but were not strictly limited to, projects that might be funded under the NHS HTA Programme.

The five categories of research are presented in *Box 2* in Chapter 2: this broad categorisation covers almost all of the measures of impact that are more selectively focused on in other studies.

### Knowledge production

The knowledge produced by research is the first output and is contained in various publications and patent applications. Any types of publications can be considered, but it is generally thought that peer-reviewed articles are the most important, and at least for biomedical research in industrialised countries it is thought reasonable to assume that the overall output of research publications is fairly represented by peer-reviewed papers in international journals. In addition to counting the number of publications, their quality and their impact can be assessed in various ways. The quality of knowledge production has traditionally been assessed by peer review, but various other methods can be applied. Papers that are accompanied by an editorial are often seen as

being of particular significance. For those studies that are included in a systematic review, there are now formal quality assessment techniques, as there are for reviews appearing in an overview.

Citation analysis can be applied to assess the impact that the specific article is having within the research community. Previous experience suggests that knowledge production will be particularly important for basic research, and certainly, on average, papers in basic research journals tend to be cited more frequently than those in clinical journals.

A journal's 'impact factor' is based on the average number of times an article in the journal is cited; it can provide a short-hand version of citation analysis by giving some indication of the importance of the journal in which an article appears. The use of impact factors in analysis of biomedical research has been criticised but, provided that care is taken, it has been shown to be of some value.

Particularly when considering research that might be aimed at potential users outside the research community, it is often desirable to use a range of publication outlets including those journals with the highest readership among the groups at whom the research is targeted. In some fields these might well be journals that do not have an impact factor but are, nevertheless, significant as vehicles for dissemination of the knowledge produced.

### Research targeting, capacity building and absorption

The better targeting of future research is frequently a key benefit from research, especially from research that is more basic and/or methodologically oriented. An indication of this comes from citation analysis. The enhanced targeting can be of the research conducted both by others and by the original researcher(s). Where follow-on research, especially by members of the original research team, is clearly associated with the original research, it can be useful to obtain information on the source and amount of such funding. Research training can be provided both as a result of the employment of staff on research projects and programmes and through explicit

funding for research training and career development. One measure of research training, which may appear crude but has nevertheless been used in previous studies, is the number and level of higher or research degrees resulting, either totally or in part, from the research funding.

### **Informing policy and product development**

Various methods have been proposed for analysing the impact of research on health policy making. Research can be used to inform policy making in a wide range of circumstances. Policy making here is interpreted very broadly to cover not just national policies of the government. It can also cover policies made by managers at many levels within a health service, policies agreed at national or local level by groups of healthcare practitioners in the form of clinical or local guidelines and policies developed by those responsible for training/education/inspection in various forms including training packages, curricula and audit and evaluative criteria. Basic research is less likely than that from clinical researchers or allied health professionals to be used to inform policy.

At a similar level, although involving very different processes, research can also be used to inform product development. Informing policies and product development are conceptually similar in that there generally has to be some subsequent adoption of the policy, or product, before the health and economic benefits can accrue.

### **Health benefits**

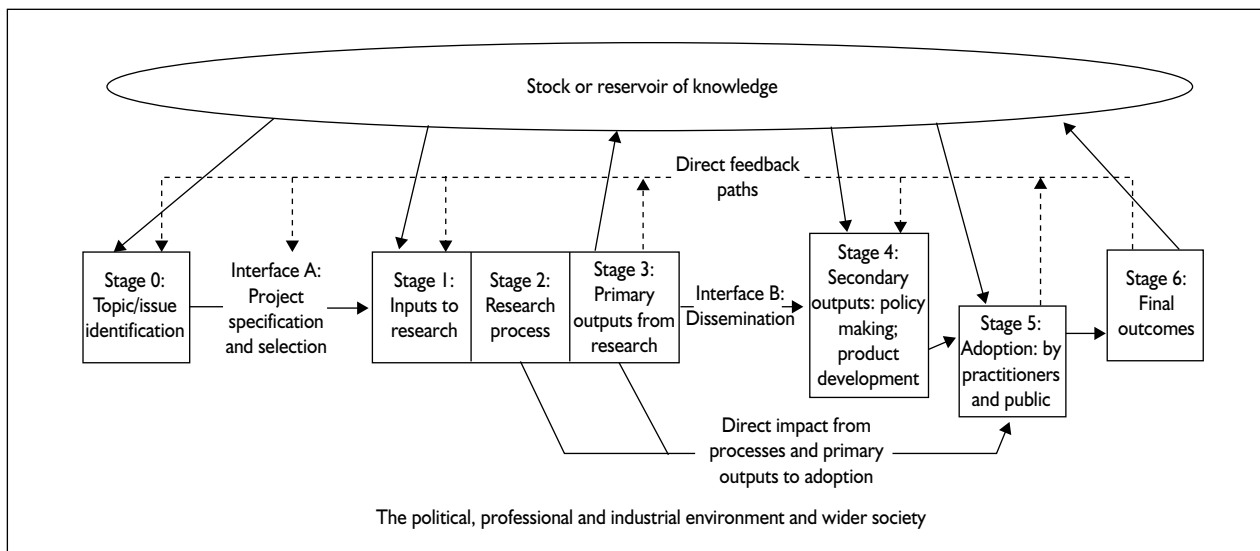
Benefits in terms of health gains might be viewed as the 'real' payback or outcomes from health research. Greater effectiveness resulting from research informed drugs or procedures will lead to increased health. Various measures of health gain exist. This category of benefits can be thought of as going wider than health gain, and some aspects can be seen as benefits to the health sector more generally. Cost savings in the provision of healthcare may result from research-informed changes in the organisation of services or in the particular therapies delivered. It might be necessary to consider various issues here. These include whether potential savings have in practice been realised – either as cash savings or as the release of resources for other valuable uses. Furthermore, it would be important to check whether, for example, costs are not simply being transferred elsewhere. Improvements could also arise in the process of healthcare delivery and these could be measured by techniques such as patient satisfaction surveys.

### **Broader economic benefits**

A range of benefits can accrue to the national economy from the commercial exploitation of research. These can take the form of employment and profits resulting from the manufacture and sale of drugs and devices. The national economy could also benefit from exports and/or import substitution.

Although there is a danger of double counting, it is probably also important to adopt a human capital approach and focus on the value of production gained from having a healthy workforce. This can be measured by examining the reduction in days off work. Typically, in a human capital approach, potential future earnings are calculated for people who, as a result of advances in medical research, can continue to contribute to national production. Those who use it, however, share the concerns that such an approach to assessing the benefits from research could have equity implications in that it would seem to favour research relevant for those of working age.

The Buxton/Hanney analytical framework of the research sequence is presented in *Figure 1*. Its main purpose is to provide a consistent conceptual structure for use in assessment. It does not attempt to model what in practice may be a very variable process. Rather, it describes a sequence of steps from identification of the topic or issue, which, in the case of the HTA Programme, is referred to as 'needs assessment'. The framework takes us through the inputs to process and primary outputs from the research project and then on to secondary outputs indicating influence on policy making through to adoption of the results by practitioners and the public, which in turn leads to final outcomes including changes in health and health service costs. Key features of the model include two important interfaces. The first is the interaction which turns an identified topic or research need into a specific project, a process involving, in the case of the HTA Programme, the interpretation of the research need as a commissioning brief, and the transformation of that by researchers into a research proposal, followed by interaction with the commissioning team who, informed by reviewers' comments, may require or request changes to the project design. The second interface is the complex process of dissemination through which the primary outputs from the research (the knowledge typically made available in academic publications and, in the case of the HTA Programme, the HTA monographs) begin to influence policy and practice.



**FIGURE 1** The Buxton/Hanney model for assessing the payback from health research. Source: Hanney and colleagues.<sup>85</sup>

A number of variations of this framework have been presented previously in the literature. The different variants reflect the extent to which the complexity (and the many elements of non-linearity) of the research sequence are included in practice in the diagrammatic representation. The model provides a consistent analytical framework to help in the assessment of impact, and is particularly useful in providing a standardised format for reporting case studies. It does not purport to capture all the complexity of research production and its influence on policy and practice.<sup>85</sup>

In the light of previous experience in applying the Buxton/Hanney payback categorisation and assessment framework, we adopted a 'mixed methods' approach, which consisted of three main elements. The first drew on information that was published or available to us from routine records at the NCCHTA. These data ranged from simple listing of the projects funded by the programme, to summary data held on file on the publications that had been reported to NCCHTA as stemming from funded projects. Some of these data are reflected in the description of the programme in Chapter 4. The second element was a questionnaire sent to all relevant principal investigators for projects funded from the inception of the HTA Programme and that had reported by June 2003. This cut-off ensured that most projects included in this survey were completed and had a potential for impact. Chapter 5 provides details of this survey. The third element was a series of 16 more detailed case studies involving desk analysis and at least one

semi-structured interview with the principal investigator or an alternative leading researcher from each project. The case studies are reported in Chapter 6.

The relationship of this project to the methods used in other recent studies is summarised in *Table 3*. It is important to note that the resources available for each study varied enormously.

The appropriateness of using a combination of these three approaches is confirmed by our analysis of the principal methods used in the studies reviewed in Chapter 2. In particular, it is clear from that review that the dilemma of the best choice between the breadth of coverage that is possible with a questionnaire approach and the depth of understanding possible with case-studies is unresolved. This choice is particularly important in that the two methods make very different demands on experienced (and hence expensive) researcher time. Case studies are resource intensive and impose organisational demands in setting up. In addition, to be of greatest effect a considerable amount of preparatory desk analysis is required before the main interview and in using, understanding and cross-checking the information provided. By contrast, an appropriate questionnaire survey can be fairly rapidly adapted from those previously used and the responses analysed quickly. To help inform this dilemma, the approach adopted here permits some comparison between the 16 case studies and the survey. Specifically, the analysis included a scoring system enabling scores based on the questionnaire to be compared with those informed additionally by the case study.

**TABLE 3** Methods used in recent applications of the Buxton/Hanney model

	<b>North Thames NHS R&amp;D Programme (Buxton et al., 2000<sup>74</sup>)</b>	<b>NHS Implementation Methods Programme (Hanney et al., 2003<sup>6</sup>)</b>	<b>UK Arthritis Research Campaign (Wooding et al., 2005<sup>75</sup>)</b>	<b>NHS HTA Programme</b>
Questionnaire to lead applicants	Yes (164)	Yes (36)	No	Yes (204)
Interview of lead applicants	Case studies only	Selective	Case studies only	Case studies only
Selective case studies	14	No	16	16
Bibliometric analysis	Case studies only	All	Case studies only	Case studies only
Documentary analysis	Case studies only	Limited for all	Case studies only	Case studies only
User surveys or interviews	Selective	Selective	Case studies only	No
Profile scoring	No	No	Case studies only	Case studies only

### Aspects excluded from our assessment

The approach we adopted essentially considers the impact of the NHS HTA Programme as the sum of the impacts from specific projects. This is reasonable to the extent that the main impacts on knowledge, on future research, on policy and administration and on practice and the healthcare system will stem from specific relevant studies or groups of relevant studies. However, there may be broader impacts generated by the programme as a whole. The preparatory study (see Chapter 1), touched on some of these based on views expressed by the respondents they interviewed. Two aspects may be particularly important.

The first is that it has been widely suggested that the HTA Programme raised the prominence and importance of health technology assessment research in the UK both among decision-makers at all levels in the healthcare system and amongst the academic health services research community. Specific aspects of this change in attitude might include the emphasis on cost-effectiveness being considered alongside, and on a par with, clinical effectiveness, and the emphasis given to establishing and promoting best methodological practice for health technology assessment through the methodology stream and subsequent programme (however, as noted above, the methodology stream was excluded from the remit of this project).

The second aspect is the perception that the programme may have raised the profile of the UK as a leading actor in this field of HTA. There has been considerable international interest in the programme as a whole, in the way it has been

organised to be actively ‘needs-led’, and the way the projects have been disseminated through the HTA monograph series. We note here these (reasonable) claims but we have not sought further specific evidence of these.

A related issue is that we have not systematically attempted to establish the impact of the HTA Programme research on policy or practice outside the UK. Generally, we would only know of such impacts if respondents to our survey or interviews were themselves aware of such impacts and brought them to our attention. Searches for citations on ISI’s Web of Knowledge and Google Scholar have identified some international impacts that have been examined.

The only quantitative indicators we have of the potential importance of the programme outside the UK are the origin of website hits and downloads of HTA monographs (see Chapter 4) and the international character of the other published outputs from the research.

Although these omissions may tend to lead us to understate the overall impact of the NHS HTA Programme, the fundamental aim of the programme was to improve the knowledge base for NHS policy and practice, not to promote a particular aspect of health-related research or to influence practice overseas.

### Ethics approval

Ethical approval was obtained from the London Multi-Centre Research Ethics Committee for both the questionnaire and the interview-based case studies.

# Chapter 4

## The NHS HTA Programme

### Introduction

This chapter briefly outlines the work of the NHS HTA Programme, including its routine arrangements for monitoring the impact of its funded research projects.

### Structure

The HTA programme is the largest programme within the National Institute for Health Research (NIHR), formerly the Department of Health R&D division. Its brief is to commission research on health technology assessment, a task which has been subcontracted since 1996 to the National Coordinating Centre for HTA (NCCHTA), based in the Wessex Institute for Health Research and Development (WIHRD) at Southampton University. NCCHTA has a three-stage process:

- Identification and prioritisation, via listing topics from a wide variety of sources and prioritising them with three panels of experts. Over 1000 topics are considered each year, drawn from a wide variety of sources, ranging from NHS organisations to an open-access web page. Systematic reviews largely by the Cochrane Collaboration make up the largest single source of topics. Around 5–10% of topics are prioritised for fuller exploration and around 1–3% lead to commissioned work (figures based on analysis of 2005; slightly different figures may apply to other years but no routine data are available). Three advisory panels of NHS experts play a key role (three over the period to 2003, since increased to four), supplemented by several hundred others who referee and comment on proposals. The criteria for prioritising projects are that the topics are of importance to the NHS, lack a commercial sponsor and are capable of yielding high-quality research.
- Commissioning and monitoring, whereby projects are commissioned by open tender and then monitored for progress. Site visits are common for clinical trials commissioned by the HTA Programme.
- Editorial and publication, with all projects published in the HTA monograph series. The

300th monograph was published at the end of 2005.

The above description applies to the period with which the present research is concerned, that is, up to June 2003. A number of changes since then are worth noting, including themed calls for research projects and in 2006 the development of a response-mode programme for clinical trials. The research projects described above, by contrast, were prioritised by the HTA Programme rather than by researchers.

### Types of research

Two main types of research are distinguished: primary and secondary. The former refers to research which generates new data and the latter to synthesis of existing research, such as systematic reviews, meta-analysis and modelling. Once topics have been prioritised, secondary research is often commissioned to inform the decision as to whether primary research should be commissioned. Primary research usually involves randomised controlled clinical trials. The commissioning of such trials, which typically cost around £1 million (2005 prices), are usually preceded by secondary research (systematic review, modelling).

Since the establishment of NICE in 1999, the HTA Programme has also commissioned TARs on each health technology-appraised process. By April 2006, NICE had published guidance on 100 technology appraisals, each accompanied by a TAR commissioned, funded and published by the HTA Programme. The cost of these reports has increased from £40,000 per standard report in 2000 to £80,000 in 2005, and £120,000 in 2005 when the cost increased further with Full Economic Costing by universities. (The shorter Single Technology Assessments developed by NICE in 2005 are funded and commissioned in the same way as the TARs which have been renamed Multiple Technology Assessments.)

In addition to NICE, the HTA Programme has a number of other policy 'customers' for secondary research reports, the largest of which is the

National Screening Committee (NSC), which plays a similar role to NICE in relation to screening. Four reports were commissioned by the NSC and 34 reports have been considered by the NSC in relation to its recommendations on screening. (Appendix 6 shows that the NSC considered 34 HTA Programme reports in establishing its recommendations, with a further 12 reports in progress.) The NSC has described its links with the HTA Programme as follows:

“Much of the NSC’s work is planned by looking at the work of the Health Technology Assessment (HTA) Programme. The production of reports commissioned by the Population Screening Panel for the HTA Programme provides the single most important and quantitatively largest influence on the work of the NSC. The NSC sits at the end of the R&D conveyor belt, picking up research reports and appraising them to consider their implications and relevance for policy making and practice.

The NSC also acts as a customer for research, identifying priorities for research in screening and influencing the Diagnostic Technologies and Screening Panel (formerly the Population Screening Panel) and the Medical Research Council to incorporate these into their rolling programme.”

*Source:* NSC website ([www.nsc.nhs.uk](http://www.nsc.nhs.uk)), UK ‘National Screening Committee’, ‘Commissioning and Responding to Research’, accessed 29 June 2006

The topics and reports considered by the NSC are listed in Appendix 6. Other policy customers include clinical directors and linked national service frameworks within the Department of Health.

## Number of projects

To June 2006, the HTA Programme had commissioned 555 research projects, of which 377 had been completed, with 178 ongoing. Of those completed, 92 or around one-quarter were clinical trials; 72 primary research projects were ongoing, of which 59 were clinical trials. Systematic reviews make up the bulk of the other projects, both for the various policy customers discussed above and as preludes to clinical trials.

Although NCCHTA does not routinely estimate a ‘failure rate’, an estimate can be made from the data used in Chapter 5 in assessing eligibility for the survey. For the period up to June 2003, 258 projects were due to have completed. Of these, 38 were methodology, responsibility for which had moved to a separate programme.

Ten projects had been ‘discontinued’ and six more described as ‘no publication was required’. The latter two groups lacked outputs that could be captured with the payback approach. These and the 38 methodology projects were rejected from the sample frame for the survey. Three projects which had been completed but whose reports had not been accepted by NCCHTA were included. Thus for the survey the sample frame was 204, made up of the original 258 but excluding the 38 methodology projects, the 10 discontinued and the six for which no publications were required.

Much depends on the definition of failure. One could be those reports refused publication by NCCHTA, in which case the failure rate would be 1.4% [made up of 3/(258 – 38)]. A wider definition might include the 10 projects which had discontinued, making the total 13 or 6%. However, the reasons for ‘discontinued’ were various, including clinical trials that failed to start or failed to recruit. The cost of these discontinued projects was considerably less than planned as a result. Thus a 6% ‘failure rate’ would be based on a broad definition of failure and should be interpreted in this light.

## Monitoring

In addition to close monitoring of the progress of individual projects, NCCHTA also monitors the number of persons accessing its reports via

- publications of research reports
- website ‘hits’.

The number of paper-based reports sold or otherwise distributed is around 100,000. The number of website views and downloads is very considerable at over 4 million or some 20,000 per report. The two main countries of origin of hits are the UK closely followed by the USA.

In addition, free CDs are widely distributed with fully searchable copies of all HTA Programme reports. Around 13,000 CDs have been distributed, mainly at conferences. No data are available on their use.

Researchers are encouraged to publish the results of HTA Programme projects in peer-reviewed journals, with such articles often appearing well before the full findings are published in the HTA Programme monograph series. NCCHTA contracts oblige researchers to inform it of all



publications and presentations for up to 3 years after the completion of the project.

### **Assessing the impact of the HTA Programme**

Although the above constitutes a level of monitoring sufficient to run the HTA Programme, several questions can be raised about the

information required to assess the impact of the programme:

- the completeness of its data on peer review and other publications
- the extent to which commissioned projects generate qualifications or lead to further research grants
- the extent to which project findings change policy or behaviour.



# Chapter 5

## Survey of lead researchers

### Introduction

This chapter summarises the methods and results of the survey of the lead researchers.

Participants were researchers who had been funded by the NHS HTA Programme for a primary or secondary research project and had submitted a final draft report (which had not necessarily been accepted for publication) between the beginning of the HTA Programme in 1993 and 30 June 2003. The survey was carried out in mid-2005.

### Exclusions

NCCHTA provided a full list of projects whose completion date was before the end of June 2003. Methodology Reviews were excluded as these were no longer funded by the HTA Programme. Projects that had been discontinued, or which had not required a publication were also excluded. Of the 258 participants eligible in 2003, 38 were excluded as methodology, 10 as 'discontinued' and six more as 'no publication was required'. The last two groups lacked outputs that could be captured with the payback approach. Some projects that had completed but whose reports had not been accepted by NCCHTA were included, on the grounds that these could have measurable outputs. These exclusions reduced the sampling frame to 204.

### Questionnaire

The layout of the 'payback approach' questionnaire used in previous work<sup>6</sup> was made more 'user friendly' and piloted during April 2005 (see Appendix 7). Apart from a minor spelling error, no other changes were made. The pilot responses have been included in the results.

Where telephone numbers were available, the Research Fellow confirmed the location of the lead researcher, updating as appropriate. This involved speaking with the researcher, their secretarial staff or other office colleagues. Where corroboration of current location was not possible, addresses for alternative authors were obtained if available. Eighteen author amendments were made and three people agreed to forward mail.

A postal rather than an email questionnaire was used as NCCHTA considered it could not make email addresses available. Questionnaires were prepared in-house and personalised to show the lead researcher by name, address and project number. A list of publications attributed to the project by NCCHTA records was included with the questionnaire, with a request for validation and updating. Each questionnaire was accompanied by an introductory letter and a stamped, addressed return envelope. Envelopes were packed in the same order, with questionnaires and letter(s) wrapped in a stamped, addressed envelope and placed so that the university logo was on display as soon as the envelope was opened. Participants with more than one HTA-funded project received details for all their studies in a single envelope.

Where substitute authors were contacted, a second letter was included explaining they had been selected because no contact was possible for the original lead researcher. Surveys were posted out on 10 and 11 May (155 and 69, respectively) with a request for return by 31 May 2005. Participants were given the option to request an electronic version of the questionnaire by email.

A second mail out/reminder was sent out on 10 June, for return by 4 July, using an identical method. A third mail out/reminder was sent to 101 participants on 13 July, with a request for return by 29 July. Following the Expert Advisory group, non-responders were emailed in a final urge to boost responses. A total of 133 valid responses were eventually received, or 65% of those surveyed.

All data, including qualitative comments gained from the responses to the questionnaires, were entered into an Access database. The number of hardcopy requests for printed copies of the reports and of web 'hits' on the NCCHTA website were also recorded on the database.

### Results

#### Response rates by types of project

Three types of project were distinguished: primary research, secondary research and TARs for the

**TABLE 4** Response rate by type of project

By type	Non-responders	Responders	Total	Response rate
Primary research	24	38	62	0.60
Secondary research	30	47	77	0.61
NICE TARs	17	48	65	0.74
Total	71	133	204	0.65

**TABLE 5** Responders/non-responders analysis for peer publications

	Number of projects	Peer publications	Publications/project
Responders	133	190	1.43
Non-responders	71	80	1.13

NICE Appraisal Committee. Primary research was mainly randomised controlled trials (RCTs). Secondary research had mainly to do with systematic reviews, meta-analysis and modelling of cost-effectiveness. This group includes health technology assessments commissioned by outside agencies other than NICE. TARs differ in being commissioned on the basis of scoping exercises carried out by NICE, with defined specifications and timetables.

The response rates by type are shown in *Table 4*, indicating a slightly higher response rate for the NICE TARs at 74% compared with an overall average of 65%.

An analysis of the number of peer publications and hits for responders and non-responders showed higher values for the responders: 1.43 for the peer-reviewed publications per project (against 1.13 for non-responders; *Table 5*) and a mean number of 21,818 website 'hits' (in comparison with 18,694 for non-responders; *Table 6*).

The response rate was higher for projects which commenced more recently, as shown in *Table 7*. A similar analysis of publications per project showed no clear trend over time, indicating that the survey results are not biased by the differential response over time.

### Publications

The total number of publications in the NCCHTA database for responders, shown in *Table 8*, was 492, made up mainly of peer-reviewed journal articles at 190 or 39% of the total. The only other large group was published presentations at 240, with the others all under 40. With published

presentations, there may be an overlap with peer-reviewed journal articles. In terms of both numbers and importance, peer-reviewed journal articles are the main category.

The gain in publications from the survey compared with NCCHTA records is also shown in *Table 8*, as 38% for peer-reviewed journal articles. The gains were considerable for the other categories but varied, as might be expected given the small numbers involved. This suggests that the methods currently used by the HTA

**TABLE 6** Responders/non-responders analysis for website 'hits'

	Number of projects	Hits (mean)
Responders	133	21,818
Non-responders	71	18,694

**TABLE 7** Response rate by time

Year	Response rate by year of project
1993	65
1994	52
1995	52
1996	59
1997	56
1998	67
1999	63
2000	74
2001	70
2002	–
2003	–
Total	65

**TABLE 8** Total numbers of publications<sup>a</sup>

	No. of projects	Peer-reviewed journal articles	Journal editorials	Journal letters	Published presentations	Citations	Published abstracts	Book	Chapter	Total
Primary research	38	85	3	0	164	3	20	1	0	
Secondary research	47	77	4	5	49	9	9	0	1	
TAR	48	28	0	4	27	3	0	0	0	
Total	133	190	7	9	240	15	29	1	1	492
Survey gain compared with NCCHTA records										
No.		73	1	3			1	1	3	
%		38	14	33			3.5	100	300	

<sup>a</sup> Excludes monographs.

**TABLE 9** Total publications for responders (including monographs)

	No. of publications
Total peer-reviewed publications	263
Publications/project	1.98
Monographs	127
Total publications including monographs	390
Total publications/project	2.93

Programme miss around one-quarter of the most important publications. (NCCHTA requires project leads to inform them of publications for up to 3 years after the end of the project. Although some journal publications precede the monograph, others are published well after the end of the project. The data collected in the survey were compared with those held by NCCHTA to see what if any gain resulted from the survey.)

The mean numbers of peer-reviewed journal articles by project and by type of project are shown in *Tables 9* and *10*, indicating 1.98 per project, rising to 2.93 when monographs (peer reviewed) are included. Higher ratios applied to primary research (3.82) and secondary research (3.36) compared with TARs (1.81).

### Qualifications gained

Twenty-eight projects reported that qualifications had been gained, leading to 31 qualifications, as shown in *Table 11*. The bulk of these (19) were associated with primary research projects.

**TABLE 10** Total peer publications per project by type of research

	Publications/project	Total publications/project
Primary research	2.92	3.82
Secondary research	2.4	3.36
TAR	0.81	1.81
Total	1.98	2.93

**TABLE 11** Qualifications gained or expected from involvement in projects

	Projects reporting qualifications	Qualifications (number)
Primary research	19	19
Secondary research	8	8
NICE TARs	1	4
Total	28	31

### Projects leading to further research funding

As shown in *Table 12*, 61 projects (46%) went on to receive further funding, with lower figures for NICE TARs, of which 32 projects (24%) stated the amount received. The total amount was £5.8 million or just under £200,000 per additional project funded. This additional funding of almost £6 million was split fairly evenly between primary and secondary research.

### Impact on policy and behaviour

Respondents were asked to indicate if, in their view, their project had impacted on policy or

**TABLE 12** Number of projects reporting further research funding<sup>a</sup>

	No of projects by type	No. of projects receiving funding	No. of projects stating amount received (funded project)	Funding (£)	Funds/funded project (£)
Primary research	38	27 (71%)	19 (50%)	3,077,092	161,952
Secondary research	47	19 (40%)	10 (21%)	2,695,577	269,558
NICE TARs	48	15 (31%)	3 (6%)	125,080	41,693
Total	133	61 (46%)	32 (24%)	5,897,749	184,305

<sup>a</sup> Four blanks taken as 'no'.

**TABLE 13** Opinion of lead researchers about existing and potential impact on policy and behaviour

	Policy			Behaviour		
	Already (%)	Future (%)	Combined <sup>a</sup> (%)	Already (%)	Future (%)	Combined <sup>a</sup> (%)
Primary research	25 (66)	27 (71)	29 (76)	17 (45)	21 (55)	23 (61)
Secondary research	27 (57)	27 (57)	36 (77)	10 (21)	22 (47)	25 (53)
NICE TARs	46 (96)	29 (60)	48 (100)	29 (60)	28 (58)	37 (77)
Total	97 (73%)	82 (62%)	113 (85%)	56 (42%)	70 (53%)	85 (64%)

<sup>a</sup> Combined = number in 'already' + number with no entry under 'already' claiming a future impact.

**TABLE 14** Conference or workshop presentations reported

	Academic audiences	Practitioner audiences	User audiences	Totals	Presentations/project
Primary research	177	158	37	372	9.8
Secondary research	154	78	12	244	5.2
NICE TARs	55	23	3	81	1.7
Total	386 (55%)	259 (37%)	52 (8%)	697 (100%)	5.2

behaviour, both to date and in the future. Some 73% of respondents claimed that their project had impacted on policy to date and 56% on behaviour (Table 13). The totals were higher for NICE TARs at 96%, as might be expected given their role. Similar figures applied to expected future impact, slightly lower compared with the past for policy and slightly higher for behaviour. When past and future impacts were combined (excluding double counting), 85% of projects claimed an impact on policy and 64% on behaviour.

### Presentations of results to different audiences

A total of 697 presentations were made, or 5.2 per project (697/133) (Table 14). Of these the bulk, 55%, were to academic audiences, followed by those to practitioners. Relatively few presentations (52 or 8%) were made to service users.

Primary research generated the highest ratio of presentations per project at 9.8, followed by

secondary research at 5.2, with NICE TARs at 1.7.

Asked if any specific one of these presentations had been of particular importance to achieving utilisation of the projects' findings, other than NICE TAR, very few responded affirmatively.

### Analysis of qualitative data

Many respondents provided additional qualitative responses, which were reviewed to identify key themes and then systematically collated and analysed. Major issues included the nature of the impact on policymaking and the bodies involved, the level of impact on behaviour, the factors associated with making an impact and difficulties with the questionnaire itself.

As noted in Table 13, a lower proportion of lead researchers from primary and secondary projects claimed that their project had made an impact on policy than did the NICE TAR respondents.

Nevertheless, for at least 33 of the 51 cases where the respondent for primary or secondary research had claimed some impact on policy, their responses indicated that this was impact at the level of national professional bodies or national agencies such as NICE, NSC, National Service Frameworks, Department of Health and/or their Scottish equivalents, or in a few instances at an international level. Within the UK, NICE was the policy-making body most frequently quoted by such respondents – at least 17 times. Many of these examples of impact on NICE may have had less significant impacts than the NICE TARs, for example, being cited in a NICE guideline. In at least eight further cases there the probability of some impact was indicated, for example resulting from ongoing discussions at a national or professional level to consider the issue.

As shown in Appendix 6, much of the HTA evidence used by the NSC has been to help inform decisions not to introduce particular screening programmes. Although important, this is sometimes difficult for researchers to record as an impact. Therefore, although at least seven examples of impact on the NSC were described in the case studies, the figure is lower than that in Appendix 6, not only because of non-responders, but also probably because some responders might have found it difficult to identify impact.

As also noted in *Table 13*, the NICE TAR respondents overwhelmingly ticked the Yes box, indicating their TAR had impacted on policy. Virtually all went on to add points about the TAR informing NICE, even in at least one case where findings were considered but the appraisal was not consistent with them.

In relation to practice, some respondents found that it was impossible to know about the level of impact, or they only knew about local examples. Many NICE TAR principal investigators thought that there was an impact on behaviour at the national level, especially since NICE guidance is mandatory. A number of others, however, said that they did not know about how far the guidance was impacting on practice. In some cases, respondents pointed out that clinicians wanted to change but had been held back until the NICE appraisal.

The timeliness and quality of the research, and liaison with stakeholders, were cited by some respondents as reasons for the impact of their work. Various NICE TAR respondents, and a few others, referred to the importance of producing findings for a policy customer, usually NICE, of

course. Some respondents, especially of systematic reviews, emphasised the importance of their study in identifying the need for further research such as RCTs. Some researchers thought that the high number of downloads received by their HTA monograph was important.

In the textual comments, the most common factor associated with a lack of impact was that it was too soon for the report to have had impact. Critical comments were made about the time taken for various aspects of the HTA process to be completed. Two respondents referred to the difficulties due to their findings being contrary to current government policy, or government policy changing. Two others referred to the problems arising when negative findings were produced. The extent to which these perceptions are well founded has not been addressed.

Finally, there were various comments on the questionnaire itself which are relevant for our analysis of the best methods for assessing research impact. The timing of questionnaires is always a problem and, in addition to those respondents who thought it was too soon to be able to answer, one respondent said the questionnaire was far too late. Furthermore, one respondent who completed several questionnaires referred to the questionnaire being unwieldy – a particular problem when a principal investigator has to complete several questionnaires. All these considerations might strengthen the case for some type of rolling system of questionnaires administered perhaps 3 or 4 years after the completion of specific projects. In addition to those respondents who said it was too early to tell whether there had been any impact, others also pointed out that it was very difficult for researchers to know about the level of impact, especially on behaviour. This is a fair point, but does not undermine the case for asking the question as a way of getting a broad-brush response from as many principal investigators as possible and as a starting point for more detailed case studies. Some correctly pointed out that the questionnaire did not allow for situations where the impact of research was to confirm existing behaviour.

### **The role of the NHS HTA Programme in enhancing utilisation**

Respondents were asked how far association with the NHS HTA Programme enhanced the utilisation of their project. A few respondents said that they did not know or it was not applicable, but of the 92 who did respond to one of the

**TABLE 15** How far association with the HTA Programme enhanced the utilisation of the project

	Not at all	A little	Moderately	Considerably	Extensively	Total
Primary research	11 (38%)	1 (4%)	6 (21%)	9 (31%)	2 (7%)	29
Secondary research	8 (25%)	0 (0%)	11 (34%)	11 (34%)	2 (6%)	32
NICE TARs	2 (6%)	0 (0%)	4 (13%)	17 (55%)	8 (26%)	31
Total	21 (23%)	1 (1%)	21 (23%)	37 (40%)	12 (13%)	92

options, 23% thought that it was not at all a help but 53% thought association with the HTA Programme considerably or extensively enhanced the utilisation of their project. The full details broken down by type of research are shown in *Table 15*. It reveals that those conducting NICE TARs were the most positive about the role of the HTA Programme in enhancing utilisation, with 81% thinking that association with the HTA Programme considerably or extensively enhanced the utilisation of their project.

## Conclusions

These results indicate that data pertinent to payback exist and can be collected. However, over one-third of projects did not respond, despite repeated reminders. Against this, a 100% response was not feasible as over the 10-year period several lead researchers had died, moved to another country or were otherwise not reachable. Analyses of the non-responders showed they had fewer recorded peer-reviewed publications and few website 'hits', indicating a possible bias towards

more successful projects in those who responded. The response rate was lower for 'older' projects.

Overall, the results are broadly in line with previous work on publications generated and on perceived impact, as summarised in *Table 2* (p. 35). Mean publications per project were 2.93 (or 1.98 excluding the monographs). The proportion reporting an impact on past policy was 73% and on behaviour 42%. When past and expected impact on policy were combined, however, 85% of respondents in this study claimed an impact on policy and 64% on behaviour. The measures compare well with those for other studies reported in Chapter 2.

The NCCHTA's present method of relying on researchers to provide data on publications is leading to less than complete data collection. Around one-quarter of publications in peer-reviewed journals are missed. Some of the data could be collected from the Internet, notably the peer-reviewed publications. Other data such as those on presentations and further research could only be collected via the relevant researchers.



## Chapter 6

# Impact of health technology assessments: 16 case studies

### Introduction

As observed in the literature, case studies are seen as an important method in the assessment of the impact of research programmes.<sup>23,40,86,87</sup> They are often seen as providing a more robust and informative analysis than can be obtained from questionnaires alone. This chapter contributes to the project's overall research questions in three ways. First, the cases provide detailed examples of the impact of projects funded by the NHS HTA Programme. Second, the cases provide data on factors associated with achieving different levels of impact. Third, comparisons with questionnaires allow analysis of the best way to assess the impact. This arises in several ways. The standard analysis of the resource-intensive case studies shows how much more they can reveal than questionnaires. Furthermore, having the two sets of data (from questionnaires and case studies) allows us to test a more systematic way of analysing the data. A system was developed to score each project's impact on the basis of the data available from NCCHTA and the survey and then to re-score using the additional information gathered in the case study. This meant that in addition to using the literature review to reveal the best methods to adopt to assess impact, we were conducting primary research into developing and testing methods to use to assess the impact of research.

### Methods

The selection of the cases to be included is a crucial issue in any analysis of a sub-set of examples chosen for detailed case study analysis. In this instance, we pioneered a stratified random selection approach to identify the 16 projects that were selected to be case studies. This is set out in more detail in a separate section following this general account of the methods used in conducting the case studies.

Prior to the interviewing in the case studies, each project was scored using data already available from NCCHTA (monographs, website 'hits', projects from NICE or NSC – Chapter 4), from

the survey, and from the impact factor and citation data for each project using the ISI Web of Knowledge (full details of the scoring methods and results are given in Appendix 9).

Then, during the case studies, additional data were collected from several sources:

- Interviews with the principal investigators, the schedules of which were based on the HERG payback framework. Prior to the interview, the interviewer read the monograph, where available, and at least one publication for the majority of interviews. The interviews were recorded and transcribed.
- Where possible, analysis of any documents referred to by the interviewee as demonstrating the impact of the HTA project, including checking for examples of other evidence that might have influenced the relevant decisions.
- Where possible, analysis of the relevant section of various databases including those of Cochrane, Clinical Evidence, NICE, SIGN and Prodigy.
- Where feasible, analysis of papers citing the main paper from the study to check for any further reviews, guidelines or other potentially important citations.
- Where the project did inform NHS guidance, especially NICE guidance, a review of the various studies of the impact of NICE (e.g. Abacus International,<sup>88</sup> Sheldon and colleagues<sup>89</sup>) to check the impact from the specific guidance in question.

When analysing the various citing documents and articles, an attempt was made to address the question of how far the particular project in question was a key piece of evidence. This was undertaken in several ways, including examining the quality ratings given to the RCTs in some of the reviews and use of the understandings gained in previous payback assessments<sup>81</sup> about how to categorise the importance of a specific citation. Ideally, and as originally intended, further triangulation would have come through interviews conducted with independent key informants who might have known about the impact of the studies.

TABLE 16 Outline of case studies

Stage		
0		Needs assessment
1	Interface A	Project specification and selection
2		Inputs
3		Processes
		Primary outputs
4	Interface B	Dissemination
5		Secondary outputs: informing policy
6		Applications by practitioners and public
		Impacts or final outcomes
	Comments	

As noted in the literature review, there is some concern about potential conflicts of interest unless others are interviewed as well as the principal investigators. While resource and time constraints made this impossible, considerable desk analysis was undertaken for most case studies.

Drawing on all the above evidence, each case study was briefly written up in about 2000 words using the stages of the HERG payback model as a series of sub-headings, but with space for possible additional comments including about areas of impact not captured by the earlier methods (Table 16).

The 16 projects were then re-scored using the same scales as used for the first payback profiles scoring but based on the full range of data available after the case studies. Inter-rater reliability was assessed using a kappa analysis (see Appendix 9 for details). Then the total scores for the initial scoring and the re-scoring were compared to see whether there was any evidence of the survey respondents exaggerating the impact of their projects.

Third, the case studies were reviewed to see what conclusions could be drawn in relation to the three issues set out above: identifying examples of impact, analysing possible factors linked to the level of impact achieved and analysing the best way to assess payback.

Fourth, a more specific cross-case study analysis was undertaken in relation to some factors linked to levels of impact. For this we listed a series of points against which each case study was analysed. These points were drawn from various sources: an initial review of the cases; rescored; comments from team members and the Advisory Group; and

points identified in the general literature on research impact. The key issues identified were systematically explored in the set of case studies as possible factors linked to impact. They included:

- Whether there was a difference in terms of impact between the projects that were already being planned (and in some cases applied for) by existing teams who then took the opportunity provided by it being a health technology assessment call, and projects where the call itself stimulated the creation of a research team.
- How far achieving impact was associated with the existence of clear customers, or 'receptor' bodies, ready to receive and potentially make use of the findings. The hypothesis here, based on previous studies (e.g. by Hanney and colleagues<sup>11</sup> and Kogan and colleagues<sup>90</sup>) is that there will be a positive correlation.
- Possible correlations between projects with high impact and strong publications, and in particular (based on previous work) whether it was possible for studies to be producing impact and yet not to have produced important peer-reviewed publications or to be receiving many citations.
- Correlations between work being assessed to be high quality (for example, in systematic reviews) and making an impact.
- How far there were above average dissemination activities by principal investigators whose work made an impact. The hypothesis here, based on previous studies (for example, those by Wooding and colleagues<sup>75</sup> and Trostle and colleagues<sup>91</sup>) is that active dissemination by researchers is likely to enhance impact.

The methods used in the analysis of the best way to assess the impact of research, included an analysis of the scoring and re-scoring exercise.

Finally, a draft of each case study was sent to the main interviewee for comment. Such a step is an important part of the validation process and not just a matter of professional courtesy.<sup>85,86</sup> The case studies were amended in the light of the comments and then a final cross-case analysis was conducted. It should be noted that an inevitable consequence of the lengthy process described above is that searches were being conducted for the various case study projects over slightly different timescales. Nevertheless, consistency was aimed for as far as possible, by conducting, for example, a final check on citations and hits on the HTA monographs for all projects at the beginning of October 2006.

## Selection of case studies

In the proposal, we committed ourselves to undertake 16 case studies and to cover each of the three main types of health technology assessment projects included in our project: primary, secondary and others/TARs. We also proposed, as in previous payback assessments by HERG, to adopt a purposive selection of cases because only a minority of research projects can be expected to produce high levels of payback and therefore a random selection might miss the key projects. A purposive range of cases would help ensure that each, or most, of the payback categories across the multi-dimensional categorisation were included in at least one or two case studies. Furthermore, it can be difficult to secure the cooperation from principal investigators to conduct case studies where the project has produced little benefit beyond a report.

There are frequently criticisms of using a purposive approach because a random element is seen as more appropriate for producing credible findings. The Ethics Committee questioned the use of a purposive selection process. Although the above arguments were accepted by the Ethics Committee as being a reasonable justification for retaining a purposive selection, further discussion led to a revised stratified random approach that was accepted by the Advisory Group. The argument that important examples of studies with impact might be missed in a random selection was mitigated by several factors in this instance. First, the HTA Programme was widely viewed as having produced a range of research with substantial impacts (see Chapter 4) and, therefore, there was less danger of important work being missed altogether in a random selection. Second, the pilot study, or scoping exercise, had conducted some case studies based on purposive selection of

projects thought likely to have a high impact. Finally, the projects were stratified not only by type of research but also by an initial simple categorisation by level of impact and some from each level were chosen for interview.

In discussion with the Advisory Group, it was finally decided that the three groups should include nine primary studies, four secondary studies and three NICE TARs. Following a brief analysis (see below), each project was allocated, by group, into three broad payback categories: high, medium and low. For the secondary projects, a random selection was made of one from each group, with a second one from the medium group being randomly chosen. For the NICE TARs, one from each group was randomly selected.

The information used to categorise the projects into the three groups came mostly from the questionnaires, but also included data from the NCCHTA about the number of hits for HTA monographs and an analysis of the journal impact factors of the articles produced by the projects. Using a simple approach involving entirely mechanistic scoring each project was given a score of 2, 1 or 0 on each of a range of categories: hits for HTA monograph (e.g. 2 = average number or above; 1 = below average; 0 = no monograph published); number of journal articles; journal impact factors for articles; impact on research training and targeting; impact on policy; impact on changed behaviour. Based on their total score, the projects were placed into the high, medium or low category prior to randomisation.

After random selection, studies were excluded and replaced if: the lead researcher was involved in the impact project (one); the lead researcher was out of the country long-term and there was no-one else available with sufficient knowledge of the project to discuss it in the depth required (one); there was a written request on the questionnaire for no further contact (one). A further few projects had to be replaced either because the lead researcher could not be contacted, or declined the invitation to participate.

## Results

The final list of 16 projects is given in *Table 17* and the full case studies are provided in Appendix 8. In *Boxes 3–5*, we provide a summary version of three of the case studies that produced most impact. These are included both to provide a flavour of the nature of the detailed studies –

**TABLE 17** The 16 projects selected for case studies**Primary research**

*Case study 1:* A randomised controlled comparison of alternative strategies in stroke care (HTA study 93/03/26)

*Case study 2:* Effectiveness of counselling, cognitive behavioural therapy and GP care for depression in general practice (HTA study 93/07/66)

*Case study 3:* Randomised evaluation of alternative electrosurgical modalities to treat bladder outflow obstruction in men with benign prostatic hyperplasia (BHP) (HTA study 94/04/09)

*Case study 4:* Efficacy and cost-effectiveness of physiotherapy for children less than four years old with cerebral palsy (HTA study 94/42/06)

*Case study 5:* Randomised controlled multiple treatment comparison to provide a cost-effectiveness rationale for the selection of antimicrobial therapy in acne (HTA study 94/48/03)

*Case study 6:* 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 (HTA study 95/07/19)

*Case study 7:* Psychological treatment in the regulation of long-term hypnotic drug use (HTA study 95/30/02)

*Case study 8:* Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain (HTA study 96/40/07)

*Case study 9:* Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography (HTA study 98/16/04)

**Secondary research**

*Case study 10:* Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis (HTA study 95/29/04)

*Case study 11:* Managing the dyspeptic patient: a systematic review and modelling exercise (HTA study 96/37/01)

*Case study 12:* Systematic review and evaluation of methods of assessing urinary incontinence (HTA study 99/29/02)

*Case study 13:* A systematic review to examine the impact of psycho-educational interventions on health outcomes and costs in adults and children with asthma (HTA study 01/16/02)

**NICE TARs**

*Case study 14:* The clinical effectiveness and cost-effectiveness of riluzole for the treatment of motor neurone disease (NICE TAR 00/01/01)

*Case study 15:* A rapid and systematic review of the evidence for the clinical effectiveness and cost-effectiveness of irintecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer (NICE TAR 00/13/01)

*Case study 16:* Effectiveness and cost-effectiveness of imatinib for first-line treatment of chronic myeloid leukaemia in chronic phase: a systematic review and economic analysis (NICE TAR 02/18/01)

organised as they are around the payback framework – and to provide some information about cases that are drawn upon in various ways in the following discussion section.

**Discussion and cross-case analysis**

Based on a combination of the general review of the cases, the specific cross-case analysis and the scoring exercise, some observations can be made on each of the main questions set out in the section 'Introduction' (p. 59).

**Payback achieved by the HTA Programme**

The case studies revealed a large diversity in the levels and form of impacts and the way they arise.

To some extent, diversity would be expected, especially given a stratified random selection process and a mixed research programme. Nevertheless, diversity existed even between projects of the same type and between ones making an impact at the same level. This is clearly seen in relation to impact on policy, where there are both many types of policies that have been informed (e.g. NICE guidance and guidelines; NSC decisions; National Service Framework; SIGN guidelines; and guidelines from many other national and international bodies) and variety of ways in which they have been informed. For example, even in relation to the three examples of NICE TARs included as case studies (Cases 14, 15 and 16), there were considerable differences in the role that they seemed to play even though all three had some impact on the policy process.

**BOX 3** Case study 1: A randomised controlled comparison of alternative strategies in stroke care (HTA study 93/03/26)

Stroke is the single most expensive disorder managed in general hospitals, with a burden likely to increase. Debates about how it should best be managed in hospitals led the NHS HTA Programme to invite tenders to compare different approaches.

The study received £500,000 from the HTA Programme to conduct a prospective, single-blind RCT. Between October 1995 and March 1998, patients were recruited from a community-based stroke register. Those with severe stroke were excluded. The study had three arms: *the stroke unit* providing 24-hour care from a specialist multidisciplinary team with clear guidelines for acute care, prevention of complications, rehabilitation and secondary prevention; *the stroke team*, which involved management on general wards with specialist team support to provide stroke assessments; and *domiciliary care*, consisting of management at home under the supervision of a GP and stroke specialist with support from a specialist team and community services for a maximum of 3 months.

In their HTA monograph, the authors concluded that, "Management of stroke patients on general medical wards, even with specialist team support, cannot be recommended because of the high mortality and dependence rate. ... a role for specialist domiciliary services for acute stroke was not supported ... the stroke unit is a more cost-effective intervention than either the stroke team or home care." (Kalra and colleagues,<sup>92</sup> p. x).

The quality of the study and its importance are indicated by publications in several major journals. Two papers were published in the very high impact *Lancet*, including the main clinical paper (by Kalra and colleagues,<sup>93</sup>) which has been cited over 60 times. Two other papers were published in *Stroke*, a major specialist journal, including the cost-effectiveness paper, the importance of which was highlighted in an accompanying editorial.

The papers are cited in several systematic reviews, including some Cochrane Reviews, especially the Stroke Unit Trialists' Collaboration 'Organised inpatient (stroke unit) care for stroke'. In this it was one of only five studies for which outcome data were available for a comparison of different forms of organised stroke unit care. It was given the top grade for its methods.

The study seems to have had a considerable impact on policy at various levels. The 'National Clinical Guidelines for Stroke' from the Royal College of Physicians cite both Kalra and colleagues' paper<sup>93</sup> directly and the Cochrane Review, again noting the strength of evidence and stressing that the recommendation that patients are admitted under the care of a specialist team for their acute care and rehabilitation should be the highest priority. Guidelines in several countries also cite papers from this study, including one from SIGN. The Stroke Council of the American Heart Association recently endorsed the guidelines from Veteran Affairs/Department of Defense and these cited both *Lancet* articles as important evidence on the organisation of stroke care.<sup>94</sup>

The study showed various gains, especially reduced mortality, from the provision of care in specialist stroke units. It is therefore reasonable to suggest that following the widespread adoption of stroke units there has been a health gain. There will also have been reduced morbidity and increased patient satisfaction from the move away from care on general wards and increased provision of specialist units. The difficulty comes in relation to the counter-factual: how far would these changes have come about without the study. Policy and practice were probably moving in the direction indicated by the findings of this study, but it provided high-quality evidence that seems to have been influential in promoting the changes.

The analysis of the 16 case studies indicated that as many as 11 were thought to have made some impact on policy at least at the level of a national professional body or policy-making body. Furthermore, the detailed case studies identified some projects (e.g. case studies 1, 2 and 11) that made a substantial impact on a range of policy-making bodies and in these cases confirmed the impression from their respective questionnaires. In particular, such case studies illustrated considerable impact in terms of knowledge production and impact on policy making, but generally less progress was made in measuring health gain and wider economic benefits. Examples of impact are given in *Boxes 3–5*. As shown in some of the boxes, some of the impact is international, including impact on guidelines in the US for the treatment of stroke and of dyspepsia (see case studies 1 and 11).

However, even with such clear examples of impact, it is not always possible to be precise about the counter-factual. It is possible, for example, that some of the changes in policy and practice might have come about because of pressure from other sources. In the instances described in the boxes, however, the evidence produced by the studies is widely seen as being of high quality and influential, in some instances being the sole or main reference given to support certain points.

The examples given in the boxes also illustrate that many of the studies helped to target future research, by both the HTA Programme and the other funders. Furthermore, three of the four systematic reviews (case studies 10, 11 and 13) played an important role in helping to target further work that was seen as directly addressing issues raised in the reviews. In some cases, funding

**BOX 4** Case study 2: Effectiveness of counselling, cognitive behaviour therapy and GP care for depression in general practice (HTA study 93/07/66)

Disorders involving symptoms of anxiety and depression are prevalent in the UK. There is growing interest in the use of counselling. In response to the HTA Programme call, the team proposed a study to determine both the clinical effectiveness and cost-effectiveness of usual GP care compared with two types of brief psychological therapy: non-directive counselling and an arm based on cognitive-behaviour therapy.

The study received £427,000 and was principally a pragmatic RCT, accompanied by two additional allocation methods allowing patient preference: the option of a specific choice of treatment (preference allocation) and the option to be randomised between the psychological therapies only. It was conducted between February 1996 and November 1997 in 24 general practices in Greater Manchester and London. A total of 464 eligible patients, aged 18 years and over, were allocated to one of the psychological therapies or usual GP care for depressive symptoms. The interventions consisted of brief psychological therapy (12 sessions maximum) or usual GP care. The patients underwent follow-up assessments at 4 and 12 months.

The authors concluded that, "In the primary care setting, non-directive counselling and cognitive-behaviour therapy were both significantly more effective clinically than usual GP care in the short term. However, there were no differences between these three treatments in either clinical outcomes or costs at the 12-month follow-up. Psychological therapy provided in primary care was found to be a cost-effective method of reducing depressive symptoms in the short term, but the comparative benefits were relatively circumscribed and did not endure over the long term. Compared with usual GP care, no differences in overall costs were observed. The additional costs associated with providing practice-based psychological therapy were recouped due to savings in visits to primary care, psychotropic medication and other specialist mental health treatments." (King and colleagues,<sup>95</sup> p. iv).

The HTA monograph received over 20,000 downloads and has been cited 23 times, and there have been a series of articles, the most important of which are linked papers in the *BMJ* giving the main clinical effectiveness and cost-effectiveness findings. The first (by Ward and colleagues<sup>96</sup>) has been cited over 70 times, and was runner-up for the Royal College of General Practitioners award for the paper with most impact on the primary care community. The project helped to target future research with team members subsequently conducting a systematic review for the HTA Programme of the effects of participants' and professionals' preferences in RCTs.

The main paper was cited in several Cochrane reviews, Clinical Evidence, *the National Service Framework for Older People*; the Department of Health's Evidence Based Clinical Practice Guideline: *Treatment Choice in Psychological Therapies and Counselling*; the NICE Clinical Guideline 23: *Depression – Management of Depression in Primary and Secondary Care*. At one point the NSF refers to the paper by Ward<sup>96</sup> as the sole evidence to support the statement: "Counselling in primary care may also be effective for depression at the less severe end of the spectrum" (Department of Health,<sup>97</sup> para. 7.27.)

There has been an increase in counselling but it is difficult to know exactly what would have happened without the study because the trend was already towards greater use of counselling.

The benefits to any one individual could be fairly large and not only could there be a considerable health gain because of the numbers involved, but also there could be a benefit to the economy in terms of a significant number of days off work avoided.

was provided by other bodies to conduct further analysis on the original dataset. Particularly important examples include case studies 6 and 7. In the former, the Department of Health provided £175,000 for a study on teenage parenthood that used the original dataset and further research with the participants and has, in turn, fed into policy documents.

Some striking examples of payback have been identified in the case studies which, for the first time as far as we could identify, were randomly selected (stratified random selection). Because impact tends to be highly skewed, examples of high impact could be missed in a random approach, which is why previous studies have adopted either a whole population or purposive selection procedure.

### Factors associated with levels of impact achieved

Of the factors considered in the cross-case analysis described above, the first two (whether the project had been prepared prior to the HTA Programme call, and whether there was a receptor body ready to use the findings) are described together here. Most, although not all, high-impact projects displayed one of these features, but not both. It is mostly for primary research that teams were already preparing proposals and mostly for secondary research, especially TARs, that a receptor body was ready to receive findings. The finding about the existence of a receptor body (especially NICE and the NSC) increasing the chances of impact is in line with the hypothesis previously developed by team members and others. The fact that the review conducted at the

**BOX 5** Case study 11: Managing the dyspeptic patient: a systematic review and modelling exercise (HTA study 96/37/01)

It was estimated that managing dyspepsia costs the NHS in excess of £500 million per year; 2% of the population consulted their GP with dyspepsia each year, and 450,000 endoscopies were performed at a cost of £90 million. Most patients undergoing endoscopy have no significant abnormality and are termed as having non-ulcer dyspepsia (NUD). The HTA vignette asked for a review of clinical effectiveness and cost-effectiveness of managing the dyspeptic patient. The study aimed to link systematic reviews with simulation modelling to provide the best available evidence for managing patients. The main questions asked in relation to the management of uninvestigated dyspepsia in primary care included assessing the effectiveness of initial pharmacological therapy, early endoscopy, *Helicobacter pylori* screening before endoscopy in patients with dyspepsia and *H. pylori* screening before eradication therapy in patients with dyspepsia.

The team received £69,000 from the HTA Programme to conduct a review from June 1998 to July 1999. It broadly followed a standard Cochrane Review approach, but with the addition of modelling of the data coming from the review. In their conclusion, the authors stated: "There is still much uncertainty around the management of dyspepsia, both uninvestigated dyspepsia and proven NUD. This review indicates that the treatment for NUD, for which the evidence is most reliable, is *H. pylori* eradication. The effect is small but cost-effective, as the treatment is potentially curative rather than just suppressive. Whether the effect is due to treating latent peptic ulcer disease or some other mechanism, the implication is that patients diagnosed on the basis of a negative endoscopy will benefit from *H. pylori* eradication" (Delaney and colleagues,<sup>98</sup> p. v).

In addition to the rapidly published HTA monograph, which has received over 20,000 downloads and 20 citations, a large number of publications arose from this study and also in some cases from the stream of work linked to it. The *BMJ* article (by Moayyedi and colleagues<sup>99</sup>), which was accompanied by an editorial, has been cited at least 148 times and, at a minimum, has clearly been taken notice of by a wide range of researchers internationally. Several Cochrane Reviews were produced by the team who formed a Cochrane Review Group. These were updated at various times as part of the ongoing work. Team members were successful in obtaining £800,000 from the Medical Research Council for an RCT to address gaps in the evidence identified by the review.

Some of the findings were directly disseminated to many practitioners in the UK, and the public, through the principal investigator being invited to be the author of the section on *H. pylori* infection in *Clinical Evidence* and of a similar section in *Best Treatments*.

The principal investigator became the technical lead for the group developing the NICE guidelines, *Dyspepsia: managing dyspepsia in adults in primary care*, and that document states: "The evidence base was derived from published reports, whose review methods are reported comprehensively" (NICE,<sup>100</sup> p. 34). The two references given for this are the HTA report and the 2000 Cochrane Review by Moayyedi and colleagues.<sup>99</sup> The details of the guidelines therefore draw substantially on the HTA Programme project and the additional work. Similarly, publications from this stream of work are well represented in the guideline from SIGN and recent official recommendations of the American Gastroenterologists Association on dyspepsia.

This impact assessment identified very little evidence about the application of the findings by practitioners and the public, although team members certainly produced papers/summaries in publications known to be targeted at practitioners and the public in both the UK and the USA. Again, little evidence was identified about the final outcomes, although the evidence has become even stronger about *H. pylori* eradication being cost-effective whereas endoscopy is not.

request of the NSC (case study 10) and all three NICE TARs did impact on policy making, whereas some of the other projects did not, is consistent with the broader pattern shown in the survey.

All three projects where the proposal had been prepared prior to the HTA Programme call had a high level of impact. Three others had the proposal to some degree developed prior to the HTA Programme call, but two of these had finished too recently to have had much impact and the third had moderate impact. Clearly, these figures are too low to be able to draw any firm conclusions, but this could be an issue to examine in further studies. Given the 'needs-led' character of the HTA Programme, it might suggest that the HTA Programme is identifying needs that others in the system too are beginning to identify.

More broadly, the case studies show detailed examples of the advantages of being 'needs-led' (e.g. funding work that might otherwise not have been funded; also providing a focus on issues of importance to the healthcare system). There are also some potential disadvantages if the needs assessment and commissioning are not done carefully; for example, in at least one instance (case study 6), the policy agenda meant that the researchers were encouraged by the HTA Programme to add an arm to their trial that had not been properly piloted and eventually this weakened the whole study. Impact is sometimes associated with researchers who felt that the commissioning process had been successful in identifying an important topic and yet sometimes allowing the researcher to focus on/add elements they felt were important (e.g. case 11). Several

projects referred to acknowledging HTA Programme funding as being important in adding credibility to their work due to being independent of drugs companies. In several cases, researchers felt it would have been difficult to get other funding for such major trials in their field (e.g. case studies 2 and 5). The HTA Programme was also reported by researchers as being a funder that aimed at generating credible needs-led research in various ways. For example, in case study 6 extra funding was provided for interpreters to ensure that the sample of disadvantaged mothers was as inclusive as possible.

Correlations between publications and impact were considered. Apart from the NICE TARs, the projects with high impact all had high scores for publications, with one exception. This finding runs somewhat contrary to some previous analyses (see, for example, Hanney and colleagues<sup>101</sup>). It is noticeable how many of the case study projects that are making most impact had their main articles published in the *Lancet* or *BMJ*.

There was not enough evidence to draw links between impact and the quality of work as judged in systematic reviews and/or guidelines: only case studies 1 and 2 produced clear evidence to support this, but here the impact of the high-quality RCT is compatible with previous work showing that high-quality research was more likely to make an impact.<sup>23,102</sup>

All high-impact projects seemed to have had active dissemination, with the possible exception of the NICE TARs, for which the direct access to the main receptor was crucial. Active dissemination (apart from TARs) may be a necessary but not sufficient condition for impact; some projects have been actively disseminated but with, as yet at least, limited impact (although in some instances such as case studies 8 and 12 this is likely to be because it is too early to expect much impact). The case studies revealed in detail how the active dissemination could take various forms. In some instances, researchers played, or are playing, an important role in the bodies formulating recommendations. These include researchers from case studies 5, 7, 10 and 11. The details of wide media coverage were discussed in several case studies, including 5 and 8.

There was some dissatisfaction with the timescales for HTA Programme work which may have led to some projects making less impact. Some felt that the commissioning process took too long, and several felt that the reviewing of their HTA report

took far too long (e.g. case studies 3, 4 and 12), although some of these also admitted that they were responsible for some of the delays. Several recommendations for improvements were made, such as ensuring the methods were seen as being appropriate from the start and, as has now already been introduced, attempting to identify possible reviewers at an early stage. Some considered that by the time the report was produced the technologies and debates had moved on (case studies 3 and 9).

Finally, some points were made in several interviews but not recorded in specific case studies. These include: a general surprise at the number of hits received by the HTA monograph, and a possible lesson for the HTA Programme is to do more to promote to authors the value of HTA reports as being widely read; several principal investigators, including some critical of delays, were pleased with support from NCCHTA, even though this was not something that was asked in the interview.

### **The best ways to assess the impact from research**

The literature review highlighted several lessons that it was hoped could be addressed in our study. In particular, there were conflicting pressures in terms of the level of resources that it was appropriate for the impact studies to employ. There was a desire to gather sufficient evidence to address issues such as the counter-factual, that is, the degree of impact that the particular study in question can have made when there were other studies and factors that might have produced the impacts. But also, there was a hope that assessment methods could be found that were less resource intensive and that could therefore be widely applied.

Several lessons could be drawn from the case studies about the appropriate methodology to use to assess impacts. The case studies provided greater detail on context and the nature of the impact than was available from the questionnaires. This took various forms, including identifying details of a wider range of impacts from some studies that were already reporting impact in the questionnaires and providing some evidence about the importance of the particular study to the policies made. Case study 10 provides a good example of the various levels of data that can be identified: important information was available from databases even before the questionnaire was completed; the questionnaire provided additional information; and the case study produced evidence of both a wider impact, for example on



guidelines from professional bodies at a European level, and about just how strong the influence of the HTA Programme study had been on decisions in the UK.

The case studies were organised according to the stages of the payback model, which not only facilitated consistency in structure and presentation but was also particularly useful in focusing on the role of the HTA Programme as a 'needs-led' programme. Furthermore, in some instances the multi-method approach adopted for the case studies facilitated the incorporation of data from previous analyses, for example ones conducted to assess the impact of NICE Guidance.<sup>88</sup> Nevertheless, in general, the case studies were still insufficiently detailed to produce much evidence about the impact of the projects on clinicians' behaviour, let alone health gain.

Undertaking the case studies following the survey facilitated the scoring and re-scoring of projects. The scoring was conducted separately in each round by two researchers and the assessment showed the inter-rater reliability of the scoring in each round to be high overall (see Appendix 9). Building on this assessment of the reliability of the scoring, further analysis indicated that in general the more detailed evidence gathered from the case studies suggested that most researchers were not making exaggerated claims in their questionnaire responses. This, in turn, suggests that questionnaires can be used to obtain a broad-brush picture of the impact of the projects, especially if attention is given to the various detailed recommendations made as to how to further improve the scoring and, in places, the wording of the questionnaire (see Appendix 9).

If the questionnaires provide a reasonably reliable source of data about the portfolio of projects, this suggests options for recommendations about case studies. Possibly a large number of brief case studies based on single interviews with principal investigators could satisfactorily supplement the data from a survey. Alternatively, if it is thought desirable to use the full multi-dimensional categorisation of paybacks, then, safe in the knowledge that the survey would provide the broad-brush data that would allow the comparisons between types of research, fewer more detailed case studies could be undertaken. Such ideas could be of particular importance as we move towards considering the role of continuous monitoring of impact.

The question of when best to make an impact assessment was dramatically illustrated by several case studies. Given the considerable elapse of time between the completion of the questionnaires and the final clearance of the case studies, there was ample scope for some studies to have produced much more in terms of output and impacts by the end of the study than the beginning. Case studies 5 and 8 provide clear examples of this.

## Summary and conclusions

Sixteen case studies were undertaken with the aim of providing more detailed examples of impact, data on the factors associated with impact and methodological analysis on the best way to assess impact. Nine primary studies, four secondary studies and three NICE TARs were selected to be case studies on the basis of stratified random selection, possibly the first time such an approach has been used.

The impact made by the 16 projects was scored on the basis of the information available from the NCCHTA database and the questionnaires. The case studies consisted of interviews with principal investigators, analysis of documents referred to by the principal investigators, analysis of key citations to the main papers and review studies of the impact of NICE. The case studies were written up using all the data available and organising it according to the stages of the payback framework.

The 16 projects were re-scored on the basis of the greater data available following the case studies. Analysis of the case studies included specific examination of a series of factors thought to be associated with achieving impact.

The case studies succeeded in their three main aims. They provided detailed examples of the impact achieved by the HTA Programme. They revealed a large diversity in the levels and forms of impacts and the way in which they arise. All the NICE TARs and more than half the remaining case study projects demonstrated some impact on policy making, at least at the level of national professional bodies or national policy-making bodies, including NICE, NSC, National Service Frameworks and the Department of Health. In particular, some case studies made considerable impact in terms of knowledge production and informing policies and confirmed the impression given by their questionnaires. They also illustrated that the projects helped to target future work by the HTA Programme and other funders.

Furthermore, some of the impact was international. Generally, less progress was made in terms of measuring health gain and wider economic benefit.

In general, the case studies provided greater understanding of how payback could be achieved, including by situating the projects in terms of wider research agendas and ongoing activities on the topic in question. Of the factors thought to be associated with achieving impact, the case studies illustrate the hypothesis proposed by team members and others that the institutional arrangements within the NHS that ensure the existence of a customer, or receptor body,<sup>11</sup> increase the chances of impact arising. A possible further hypothesis began to emerge from the analysis of the case studies and this related to the high level of impact achieved by cases where the research team had already prepared the proposal and then the needs assessment by the HTA Programme led to them issuing a call in that specific area.

More broadly, the case studies show detailed advantages of being 'needs-led' (e.g. funding work that might otherwise not have been funded, and also providing a focus on issues of importance to the healthcare system). The case studies provide at least one example, however, where difficulties arose because changes in design were imposed for policy reasons.

Apart from the NICE TARs, all the case studies with high impact were associated with successful

publications, with one exception. Similarly, they were all associated with active dissemination, again with the possible exception of the NICE TARs, where there was very important – but possibly narrow – dissemination. However, other projects too were actively disseminated, so it might be a necessary element of achieving impact but not sufficient in itself. Although there was broad satisfaction with the support from NCCHTA, there was some dissatisfaction with various aspects of the sometimes long timescales involved.

Finally, lessons could be drawn about the appropriate methodology. The case studies provided more detail on the impact made and, being organised according to the stages of the payback framework, they not only facilitated consistency in structure and presentation but also were particularly useful in focusing attention on the role of the HTA Programme as a 'needs-led' programme. In general, the case studies were still insufficiently detailed to produce much measurement about the impact of the projects on the clinicians' behaviour, let alone the level of any health gain. The kappa analysis showed the inter-rater reliability of the scoring and re-scoring of projects to be reasonably high and analysis of the two rounds of scoring suggested that the further evidence gathered in the case studies, although still usually from the researchers, provided some indications that most researchers were not making exaggerated claims in their questionnaire responses.

# Chapter 7

## Summary and discussion

This chapter summarises the results of the assessment and goes on to discuss their implications.

### Summary of findings

#### Literature review (Chapter 2)

The systematic review of the literature on impact assessment identified a relevant and interesting literature of around 200 papers. From these, 46 papers were included for detailed review, to identify useful approaches. Although a small number were conceptual papers, introducing frameworks for impact assessment, the majority offered useful insights to the empirical application of the methods suggested.

The frameworks identified tended to focus on specific impact assessments of designated research programmes or groups. However, some were more generic, applied across a range of programmes (e.g. the payback approach), or could be generalised to other settings. Almost all frameworks involved one or more of the core research activities of desk analysis, questionnaires, interviews and case studies.

Limitations in the conduct of systematic reviews on topics such as this include difficulties interpreting the findings from impact studies due to the context specific nature of the studies and differences between the research programmes assessed (for example, their scope, the period of research considered, the different methods employed and ambiguities over what is meant by impact on policy). Also of note are the potential conflicts of interest in the studies reported (i.e. often undertaken by the funding body interested in identifying impact), and possible selection biases in the studies that have been subject to impact assessment.

The studies identified and discussed were largely successful in presenting quantifiable measures of impact from research funding. Some assessed this from a multi-dimensional perspective, and others targeted a single aspect (e.g. knowledge production, or policy). When a broad approach is taken to impact assessment (i.e. multiple

dimensions of impact), it is clear that such impact is more easily identified for knowledge production, through the reporting and capture of dissemination activities, than is the case for impacts on policy, practice and health outcomes. In these latter domains, impact assessment often involves judgements on the use made of research findings, which may involve assessing the contribution of the research to a broader body of evidence. It may be necessary to seek advice on such judgements from a broader sample than that usually used in impact assessment projects (including users, policy makers and decision-makers). Greater impact has been noted in studies that have focused on the impact of research on health policy (e.g. Hailey and colleagues<sup>36</sup>), using more detailed methods to seek opinion on policy impact, and relying on judgements on level of impact.

While the review presents a broad literature on approaches for impact assessment, the literature continues to develop, with recent studies published by Johnston and colleagues<sup>65</sup> on a US programme of clinical trials and Kingwell and colleagues<sup>66</sup> reporting on the assessment of impact from the Australian NHMRC-funded research.

The review presented in this report adds to the general literature on impact assessment: first, as noted above, the enormous difficulties in assessing the later categories of impact; second, and despite these difficulties, higher levels of impact on policy emerge compared with what has often been claimed. This may be due to the studies examined having taken the research conducted as their starting point.

Of particular interest to the current report is the evidence around the use of the payback approach presented by Buxton and Hanney.<sup>23</sup> This is the most commonly used approach both in the UK and elsewhere, and was identified *a priori* as the framework to guide the impact assessment process undertaken in this project for the NHS HTA Programme. The international literature review showed eight empirical applications reported up to 2005, plus other impact assessment studies drawing from the payback literature (e.g. Ferguson and colleagues<sup>32</sup>). The adherence, across these

studies, to the payback approach has varied, providing evidence of flexibility within the approach, but also showing that the use of such a conceptual approach may not always be necessary (or desirable). The literature provides some evidence of the acceptability of the payback approach to the research community, and also to policy makers. It also indicates that it can be applied to different settings. The primary tools of the payback approach can be adapted to specific assessment questions, and although it is clear that not all of the categories of payback defined by the approach are readily accessible in the empirical assessments (e.g. health sector benefits, broader economic benefits), such categories of payback are undoubtedly desirable and most commentators would agree that they should remain a target for any future assessment initiative.

In terms of informing further work, the review provided various lessons for our specific study of the impact from the NHS HTA Programme. One of the key lessons related to the complexity of finding appropriate comparators for the NHS HTA Programme. Possibly the NICE TARs from the NHS HTA Programme are broadly comparable with the health technology assessment group of studies reported in *Table 2* because they are usually programmes of assessments explicitly produced for decision-making bodies. In many ways, the primary and secondary studies from the NHS HTA Programme are more like many of the 'other health research programmes' listed in *Table 2*.

The very diversity in the existing and emerging studies in this field leads to the suggestion that a research project to collate the existing and emerging studies and analyse them in various ways would be of benefit. Finally, the feasibility of assessing the health gain and economic benefits should be considered through undertaking detailed case studies.

#### **The NHS HTA Programme (Chapter 4)**

Key characteristics of the NHS HTA Programme include its emphasis on rigorous scientific research on health technologies that matter to the NHS. Around 1000 suggested topics are examined each year and prioritised with the help of around 1000 experts, mainly NHS clinicians. The topics come from a wide variety of sources, ranging from NHS organisations to an open-access web page. Systematic reviews, largely by the Cochrane Collaboration, make up the largest single source of topics. About 6–10% of topics are prioritised for fuller exploration and around 1–3% lead to

commissioned work. Advisory panels of NHS experts play a key role, supplemented by several hundred others who referee and comment on proposals.

The HTA Programme has important policy 'customers', such as NICE and NSC. Systematic reviews, meta-analysis and cost-effectiveness modelling tend to be the 'products' for these 'customers'.

#### **Survey results summary (Chapter 5)**

These results indicate that data pertinent to payback exist and can be collected. However, over one-third of projects did not respond, despite repeated reminders. Against this, a 100% response was not feasible as several lead researchers had died, moved to another country or were otherwise not reachable.

Overall, the results showed a wide range of publications from the HTA Programme, considerable effort into dissemination and over two-thirds of projects claiming to have had an impact on policy. From the survey, it was clear that the HTA Programme often had an impact on national policies through professional associations and national agencies, including NICE, NSC and the Department of Health.

Mean peer-reviewed publications per project were 2.93 (or 1.98 excluding the monographs). The proportion reporting an impact on past policy was 73% and on behaviour 42%. When past and expected future impacts were combined, 85% of responders claimed an impact on policy and 64% on behaviour.

As might have been expected, there were variations between the impact of different types of research funded by the HTA Programme. In terms of publications, the primary and secondary research projects produced considerably more peer-reviewed publications than the NICE TARs. As suggested above, perhaps the most suitable comparators for the NICE TARs would be those HTA Programmes listed in *Table 2*. No information was recorded about the publications produced by those Programmes. Nevertheless, even including the NICE TARs, the NHS HTA Programme had a higher average number of peer-reviewed publications than the 'other health research programmes' listed in *Table 2*.

Some 96% of NICE TARs reported an impact on policy compared with 60% of the other studies. In this regard, the NICE TARs compare well with the

figures for the HTA Programmes listed in *Table 2*, where the impact on policy was high across the board. The 60% of primary and secondary projects from the NHS HTA Programme reporting an impact on policy is a higher figure than that reported for any of the 'other health research programmes'. The impact of the NICE TARs on practitioners' behaviour was higher than that of the impact from the primary and secondary research of the NHS HTA Programme, which in turn was about in the middle of the range of that of 'other health research programmes' in *Table 2*.

Several suggestions were made on how collection of payback data by NCCHTA might be improved. Publications in peer-reviewed journals might best be identified by regular literature searches, which are then fed back to researchers for clarification and additions. The archiving of documents by NCCHTA showing the evolution of projects (from brief to protocol to end report) also needs to be improved, given the difficulties encountered in obtaining such data for this report.

There was a positive response to the question about how far the project's association with the HTA Programme had enhanced the level of utilisation, with 53% of the responders ticking the 'considerably' or 'extensively' boxes.

### Case studies (Chapter 6)

Sixteen case studies provided more detailed examples of impact and the factors associated with impact. They also permitted comparison with the survey as a means of collecting data. Nine primary research studies, four secondary studies and three NICE TARs were selected as case studies on the basis of stratified random selection, probably the first time such an approach has been used.

The impact of each project was scored on the basis of the information available from the NCCHTA database and the survey questionnaires. The case studies consisted of interviews with principal investigators, analysis of documents referred to by the principal investigators, analysis of key citations to the main papers, and review of studies of the impact of NICE. The case studies were written up using all the data available and organised according to the stages of the payback framework.

The 16 projects were re-scored on the basis of the greater data available following the case studies. Analysis of the case studies included specific examination of a series of factors thought to be associated with achieving impact.

The case studies succeeded in their three main aims. They provided detailed examples of the impact achieved by the HTA Programme. They revealed considerable diversity in the levels and forms of impacts and the way in which they arise. All the NICE TARs and more than half of the other case study projects demonstrated some impact on policy making, at least at the level of national professional bodies, or national bodies including NICE, NSC and the Department of Health. Some case studies indicated considerable impact in terms of knowledge production and informing policies, confirming the impression given by their questionnaires. They also illustrated how projects helped to target future research priorities and funding by the HTA Programme and other funders. Some of the impact was international. Less progress was made in terms of measuring health gain and wider economic benefit.

The case studies provided greater understanding of how payback could be achieved, mainly by situating the projects in terms of the wider research agenda. Of the factors thought to be associated with achieving impact, the case studies support the hypothesis (proposed by team members and others) on the importance of a 'customer' or receptor body<sup>11</sup> increasing the chances of impact arising. A further hypothesis emerging from the case studies related to the higher level of impact when the research team had already developed the research ideas and study protocol prior to a proposal being tendered by the HTA Programme.

More broadly, the case studies show the advantages of the 'needs-led' HTA Programme in providing a focus on issues of importance to the healthcare system and in funding work that might otherwise not have been funded. The case studies provide one example, however, where difficulties arose because the study design was changed at a late stage.

Apart from the NICE TARs, all but one of the case studies with high impact were associated with successful publications. Similarly, all had active dissemination, again with the possible exception of the NICE TARs, which had a more specific focus. However, other projects too were actively disseminated, so it might be a necessary but not sufficient condition for achieving impact. Although there was broad satisfaction with the NCCHTA, there was some dissatisfaction with the long timescales involved.

Finally, lessons could be drawn about the appropriate methodology. The case studies provided more detail on the impact made. Being organised according to the stages of the payback framework not only facilitated consistency in structure and presentation for case studies, but was also particularly useful in focusing on the role of the HTA Programme as a ‘needs-led’ programme. In general, the case studies were still insufficiently detailed to produce much measurement about the impact of the projects on the clinicians’ behaviour, let alone the level of any health gain. The kappa analysis showed the inter-rater reliability of the scoring and re-scoring of projects to be reasonably high and further analysis of the two rounds of scoring suggested that the further evidence gathered in the case studies, although still usually from the researchers, indicated that most researchers were not making exaggerated claims in their questionnaire responses.

## Discussion

The remainder of this chapter considers the following questions:

1. What has been the payback from the HTA Programme?
2. What factors seem to be associated with ‘higher’ impact?
3. Implications for the NHS:
  - (a) Implications for UK HTA Programme policy?
  - (b) Generalisability of conclusions to other health technology assessment programmes?
4. What have we learnt about assessment of payback?
5. Scope/methods for future routine monitoring of the HTA programme?
  - (a) Implications for future research on assessment?

### What has been the payback from the HTA Programme?

This study, based on databases, questionnaires and case studies, shows that overall the HTA Programme has had considerable impact in some of the payback categories, especially in terms of knowledge production and impact on policy. The number of peer-reviewed publications per project, at almost three, was highly commendable, comparing well with other programmes. (There are also some important examples identified of where the HTA Programme studies have targeted future research by the HTA Programme itself and other funders.)

This overall result needs to be interpreted against the key characteristics of the HTA Programme discussed above as ‘needs-led’ topics and using scientific methods. The former ensures that topics are commissioned on technologies of importance to the NHS. These technologies have often been of little interest to other funders, particularly private sector (no intellectual property rights), but also public sector funders such as the MRC (which is mainly ‘science-driven’) or the medical charities (also largely ‘science-driven’). This has led to research on topics such as back pain, rehabilitation, acupuncture and counselling, and also novel comparisons (psychological interventions compared with medicine or surgery).

One might have expected a worse performance from research on such topics based on the conventional measures of contributions to knowledge as measured by peer-reviewed publications. Conversely, however, one could argue that the novelty of research on these topics might facilitate publication, but perhaps in less mainstream medical journals. The fact that the mean number of publications for the programme as a whole compares well with other programmes, and largely in leading medical journals, suggests that the HTA Programme has succeeded in establishing research in these novel topics. This success may reflect the programme’s emphasis on rigorous science, which is often path breaking in relation to the topics selected.

Another characteristic of the HTA Programme, noted above, is its direct work for policy ‘customers’ (or ‘receptor bodies’) such as NICE and NSC. One might expect such work to score highly on impact on policy and perhaps on practice. The impact on policy is clearest in relation to studies, usually secondary research, commissioned for these ‘customers’, but also applies to clinical trials, some of which have had considerable policy impact as shown by the case studies.

Although the research projects can be readily classified by primary/secondary research, this study has also revealed a large diversity in the levels and form of impacts and the way they arise. This diversity is clearly seen in relation to impact on policy where there are both many types of policies that have been informed (e.g. NICE guidance and guidelines; NSC; National Service Frameworks; guidelines from many bodies) and many ways in which they have been informed. For example, each of the three TAR case studies played different roles.

The case studies identify some projects (including case studies 1, 2 and 11) that made a substantial impact confirming the questionnaires.

However, even in clear examples of impact, it is not always possible to be precise about the counter-factual. It is possible, for example, that some of these projects may have been funded later or in slightly different forms by other funders.

Some of the impact is international, including impact on guidelines for the treatment of stroke and dyspepsia in the USA (see case studies 1 and 11).

### **What factors seem to be associated with 'higher' impact?**

Two obvious factors linked to higher levels of impact on policy are topic relevance and existence of a policy 'customer'. This study suggests that the HTA Programme appears to have been largely successful in identifying topics of relevance to the NHS and which often had policy 'customers' (such as NICE and NSC). In relation to knowledge generation, in addition to topic relevance, the quality of the research is important. This study shows that the HTA Programme has resulted in high-quality peer-reviewed publications from a large number of projects. The specified use of rigorous methods (systematic reviews, meta-analysis, RCTs) to a large degree ensured high-quality research, particularly when coupled with rigorous peer review.

Evidence from questionnaires and case studies indicates that being a 'needs-led' programme has given rise to a relatively high level of impact. For example, many health technology assessment projects have informed the work of the NSC in addition to NICE. The case studies show detailed examples of the advantages of being 'needs-led' (e.g. funding work, as above, that might otherwise not have been funded; also providing a focus on issues of importance to the healthcare system). However, in one case study the policy agenda led to the researchers being encouraged by the HTA Programme to add an additional 'arm' to their trial, which weakened the study.

Impact varied by type of project. The questionnaires and case studies confirmed what was shown in the NCCHTA data – that NICE TARs have a greater policy impact but are less likely to lead to further publications. Where there was impact from non-NICE TAR projects, often one or more of the researchers played an important role in promoting the findings. A preliminary

impression (which needs more work to confirm) suggests that projects making a big impact often had publications in important journals.

The survey and the case studies revealed some dissatisfaction with certain aspects of the programme, particularly the length of time taken to agree funding and to publish the monographs. Some researchers considered that these delays reduced the impact of the research.

The HTA Programme was seen as an important and independent funding source. Several projects referred to HTA funding as having been important in adding credibility to their work (e.g. independent of funding from the pharmaceutical industry). In several cases, researchers felt it would have been difficult to obtain alternative funding for specific clinical trials.

### **Implications for the NHS**

A major implication of the impact assessment is that given the level of impact that is being achieved, it would seem reasonable to argue that from the perspective of the NHS it might be desirable for the HTA Programme to expand. This would particularly be the case if, at the same time, NHS customers for the research become more pro-active in seeking out information and implementing it.

### **Implications for UK HTA Programme policy?**

This study suggests some implications for enhancing the HTA Programme, as follows:

- Consider how to speed up the commissioning and reviewing processes. To the extent that the delays are due to peer reviewers, consideration should be given to paying them, particularly since the contribution required can be considerable. Early identification of possible reviewers was introduced a few years ago and is something some principal investigators called for.
- Improving the collection of data on publications. The survey showed that up to one-quarter of all publications were not being collated by NCCHTA. Use of literature searches may provide an alternative to relying on researchers to inform the programme of publications
- Maintaining fuller records of the entire process whereby topics are prioritised and tendered. Although the records kept seem adequate for managing the programme, they were less than complete for monitoring and evaluating the successes and failures of the programme.

- Being careful about imposing changes on design, particularly for policy rather than scientific reasons. One case study shows the problems caused by such change.
- Sharing data on publication and Internet hits more widely, particularly with researchers who were often unaware of the frequency with which their reports were being accessed.

### **Generalisability of conclusions to other health technology assessment programmes?**

In generalising, it is necessary to be careful about comparing like with like. Many other health technology assessment programmes do not include clinical trials for instance. None provide comparable input to NICE-type decision-making bodies.

More focused comparisons could well be helpful for benchmarking. This might require comparing different elements of HTA Programme work with different agencies (for example, compare clinical trials with other funders of clinical trials, such as the MRC in the UK or with the National Institutes of Health in the USA).

Assessments of entire health technology assessment programmes are rare and poorly reported. The recent assessment by the AHFMR of evaluations of HTA Programmes found that only two of the 16 assessments considered had been reported in the peer-reviewed literature.<sup>103</sup> Our review found few assessments of entire health technology assessment programmes. Most previous evaluations have not looked comprehensively at impact.

Given the dual nature of the HTA Programme in terms of some of it being much more directly fed into policy than other parts, there might be lessons from different parts of it for different health technology assessment programmes.

### **What have we learnt about assessment of payback?**

This study has contributed to the body of knowledge on assessment of the impact of research through a literature review and primary research. The former provided a systematic analysis of a large and diverse literature from which several key messages stand out. Many of the studies identified were successful in demonstrating a range of impacts, but it is often easier to assess impact in terms of knowledge production than in relation to impact on practice and health gains. Assessing impact on policy making involves making judgements. Various studies report a

higher level of impact on policy than has often been assumed to occur. This may be due to their starting point being the research conducted, rather than the policies made. In many cases they focused on 'needs-led' research. Health research impact assessment is currently attracting considerable attention. Several key articles published since the review was completed have been noted and fed into the discussion below. Nevertheless, there have been comparatively few attempts to develop conceptual frameworks for general application. The Buxton and Hanney payback framework is the one most widely applied.

Conflicts of interest can often occur in that most impact assessments are sponsored by the funder of the research programme being assessed. Much of the information about the impacts comes from the researchers themselves or the clients for the original research.

Lessons have been learnt by this study. The use of broad data collection methods (analysis of databases and questionnaires) and more concentrated ones (case studies) enabled the general picture to be enhanced with detailed illustrative examples. This study adopted an innovative stratified random selection approach for choosing case studies which worked well and resulted in a representative sample of case studies.

This study developed work aimed at devising ways to score a project's impact. This can be based first on broad data collection, and then compared with more detailed case studies. Reasonably high levels of inter-rater reliability applied in both rounds, although higher for some scoring scales than others. Further refinements would be possible given the lessons learnt. Nevertheless, it was considered sufficiently robust to allow comparisons between the two rounds of scoring. The further evidence gathered in the case studies, although still mostly collected from the researchers, suggested that most researchers were not making exaggerated claims in their questionnaire responses. This, in turn, might confirm the usefulness of questionnaires in assessing impact.

The payback framework allows the collection and organisation of a wide range of material on the impacts of health research. It also facilitates the analysis in the case studies of the importance of the needs-led element to the achievement of payback. In applying the multi-dimensional categorisation, it proved more difficult to assess



the impact on behaviour and health outcomes. This is due to timescale and clarity of evidence. Nevertheless, substantial impacts on policy making were identified.

The timing of impact studies is difficult to get right. Some researchers thought it too early to make assessments even though these were being undertaken some time after the project had finished. Some case studies showed how rapidly impact indicators (such as web hits on a particular report) could change and how during the course of conducting and clearing a case study a project could make (additional) impacts on policy that were not in existence at the start of the case study.

### **Scope/methods for future impact assessments and routine monitoring of the HTA programme?**

It is worth considering a rolling programme of questionnaires that might merge into a regular monitoring system. Such a development could improve the timing of impact assessment; a recommended period would be about 3 years after the publication of the HTA report to allow time for impacts. To encourage a higher response rate, it might be necessary to introduce a contractual requirement to complete a post-completion impact questionnaire. These suggestions are in addition to the recommendation already made about improving the routine collection of information on publications.

Although it might be important to conduct a few detailed case studies along the lines suggested above, it might also be worth considering how far there would be scope for instituting a rolling programme of mini-case studies that would be informed by the work described in this report and chosen on the basis of the evidence from the routine monitoring. They could build into a portfolio that would be available for the HTA Programme to use to demonstrate its impact, but also address organisational and management issues about how best to run a programme to maximise the impacts achieved.

### **Implications for future research on impact assessment**

From the literature review and this impact assessment, it is possible to identify three main areas for further research:

1. a major effort to explore the feasibility of analysing the later payback categories of health gain and economic benefits from programmes of health research

2. specific issues around the implementation of the payback framework
3. an ongoing collation and analysis of the diverse and ever-increasing number of impact assessment studies.

*Area 1:* There is an overlapping debate about the desirability and feasibility of attempting to assess the impact of research on health gain and economic benefit. Broadly, there is seen to be a spectrum in which impact is more easily identified for knowledge production activities than for impact on policy and even more so than for impact on practice or outcomes such as health gain or economic benefit. The debate about the desirability of assessing the later parts of the spectrum is a matter of policy. Some argue that research funding bodies and systems should concentrate on those issues for which they have most direct responsibility, that is, the production of knowledge. Others argue that a health research system should be concerned with what the broad impact of its research has been and how it could be enhanced. From the perspective of informing this debate, it could be important to make progress in showing the feasibility of assessing health gain and economic benefit. This presents itself as a major challenge for future research.

One possible way of advancing such a stream of work would be to develop case study approaches to help assess the impact on behaviour and the gains in the health sector and the broader economic benefits. The recent article by Johnston and colleagues<sup>65</sup> used secondary sources about the uptake and benefits of certain interventions to assess the economic value of key clinical trials conducted by NINDS. In the UK context it would be necessary to undertake primary research using the type of payback approach described by Buxton and colleagues<sup>10</sup> to establish that the particular research was responsible for the introduction of the specified intervention that leads to the health and economic gains. Questions about the counterfactual would also need to be addressed. Such detailed case studies might explore the relative contributions of related projects on particular topics. Whereas the study by Johnston and colleagues<sup>65</sup> examined the impact from a set of clinical trials from NINDS, it is possible that for programmes such as the NHS HTA Programme, any attempt to use such intensive case studies would be on a more selective basis.

*Area 2:* The detailed application of the payback framework. The textual comments made in the questionnaires (see Chapter 5) identified several

issues where the wording of future questionnaires could be improved. The discussion about the role of case studies based on the payback framework showed they can facilitate analysis of contextual issues. This has highlighted several points that could be turned into items for the questionnaire, for example, how far the research team was already formed and had prepared a similar research proposal.

The scoring system could also be refined, especially where the inter-rater reliability was lowest, that is, in relation to the impact on policy and practice. Some iteration is required between developments in questionnaires and scoring: the wording of the questionnaire has to be appropriate to gather the data for scoring. The scoring of projects has uncovered problems with the wording of the questionnaire (see Appendix 9).

*Area 3:* There is considerable diversity amongst the studies included in the literature review, with a

growing body of research in this field, and as we have seen some important studies have been reported since the formal cut-off point for this review. This collection of studies is very different from that in most medical fields because usually each impact study is a unique analysis of a programme or body of research and it would be unusual for another impact study to be undertaken on the same body of research. This means that the standard systematic review techniques are not appropriate. There might, however, be considerable value in a research programme or project that collated health research impact studies in an ongoing manner and analysed them in a consistent fashion even though the studies themselves would probably continue to be diverse. One of the issues that such an ongoing impact-research collation and analysis programme could consider in more detail than has been possible here is the question of the time elapsed before impact occurs and, building on that, how it might be possible to assess impact over different timescales.



## Acknowledgements

The Advisory Group chaired by Chris Henshall is made up of Philip Home, Peter Sneddon, Kalipso Chalkidou, Isobel Walter and Lynn Kerridge. They also contributed to the project by commenting on various drafts. Ansgar Gerhardus and Wija Oortwijn also commented on the key references identified. Diane Coulson carried out the survey. Teri Jones assisted with the scoring. Luminita Grigore helped assemble the final report. Alison Price helped with the literature search. Thanks are due to all of them.

### **Contribution of authors**

Martin Buxton (Professor of Health Economics) drafted Chapters 1 and 3. Steve Hanney (Senior Research Fellow) and Colin Green (Senior Lecturer in Health Economics) drafted Chapter 2. James Raftery (Director of WIHRD) drafted Chapters 4, 5 and 7. Steve Hanney drafted Chapter 6 and the qualitative part of Chapter 5. All authors commented on various drafts of the report.





## References

1. Department of Health. National Coordinating Centre for Health Technology Assessment. *The Annual Report of the NHS Health Technology Assessment Programme*. London: Department of Health; 2000.
2. The Academy of Medical Sciences, Medical Research Council, Wellcome Trust. *Evaluation Forum. Medical research: assessing the benefits to society*. London: The Academy of Medical Sciences, Medical Research Council, Wellcome Trust; 2006.
3. Murphy K, Topel R. Measuring the gains from medical research. An economic approach. Chicago: University of Chicago Press, 2003.
4. National Audit Office. *Getting the evidence: using research in policy making*. Report HC 586-1, 2002–3. London: National Audit Office; 2003.
5. Wooding S, Hanney S, Buxton M, Grant J. *The returns from arthritis research. Volume 1: approach, analysis and recommendations*. Cambridge: RAND, Europe; 2004.
6. Hanney S, Soper B, Buxton M. *Evaluation of the NHS R&D Methods Programme*. HERG Report 29. Uxbridge: Brunel University; 2003.
7. Soper B, Hanney S. Lessons from the evaluation of the UK's NHS R&D Implementation Methods Programme. *Implementation Sci* 2007;1:2.
8. Nutley S, Percy-Smith J, Solebury W. *Models of research impact: a cross sector review of literature and practice*. London: Learning and Skills Research Centre; 2003.
9. Gerhardus A, Dintsios C. *The impact of HTA reports on health policy – a systematic review*. Cologne: German Agency of Health Technology Assessment at German Institute for Medical Documentation and Information (DAHTA) (DIMDI); 2004 (in German).
10. Buxton M, Hanney S, Jones T. Estimating the economic value to societies of the impact of health research: a critical review. *Bull World Health Organ* 2004;82:733–9.
11. Hanney SR, Gonzalez-Block MA, Buxton MJ, Kogan M. The utilisation of health research in policy-making: concepts, examples and methods of assessment. *Health Res Policy Syst* 2003;1(2).
12. Pardes H, Manton J, Lander E, Tolley H, Ullian A, Palmer H. Effects of medical research on health care and the economy. *Science* 1999;283:36–7.
13. Cronin E, Normand C. *Methodologies for evaluating R&D programmes – a literature review*. London: London School of Hygiene and Tropical Medicine; 1999.
14. CRD. *Undertaking systematic reviews of research on effectiveness*. York: Centre for Reviews and Dissemination; 2000.
15. Adamson TA. The productivity and impact of psychiatric research on development of mental health services in Nigeria. *West Afr J Med* 1998; 17:243–7.
16. Alberta Heritage Foundation for Medical Research (AHFMR). *A study of the impact of 2000–2001 HTA products*. Edmonton: AHFMR; 2002.
17. Alberta Heritage Foundation for Medical Research (AHFMR). *A study of the impact of 2001–2002 health technology assessment products: final report*. Edmonton: AHFMR; 2003.
18. Alberta Heritage Foundation for Medical Research (AHFMR). *Assessment of health research fund outputs and outcomes: 1995–2003*. Edmonton: AHFMR; 2003.
19. Bjerregaard P, Curtis T. Impact of public health research in Greenland. *Int J Circumpolar Health* 2004;63:214–20.
20. Bos M, Carlsson P, Kooj S, Liaropoulos L, Sampietro-Colom L, Schilling J, et al. *Technology assessment and coverage policy: the case of invasive cardiology therapy in five European countries*. Barcelona: Catalan Agency for Health Technology Assessment and Research; 1997.
21. Britton M, Jonsson E. Impact of health technology assessments. Some experiences of SBU. *Int J Technol Assess Health Care* 2002;18:824–31.
22. Brownson RC, Simoes EJ. Measuring the impact of prevention research on public health practice. *Am J Prev Med* 1999;16 (3 Suppl):72–9.
23. Buxton M, Hanney S. How can payback from health services research be assessed?. *J Health Serv Res Policy* 1996;1:35–43.
24. Buxton M, Hanney S, Packwood T, Roberts S. *Assessing the benefits from North Thames R&D*. HERG Report 25. Uxbridge: Brunel University; 1999.
25. Buxton M, Schneider W. *Assessing the 'payback' from AHFMR-funded research*. Edmonton: AHFMR; 1999.
26. COHRED Working Group on Research to Action and Policy. *Lessons in research to action and policy – case studies from seven countries*. Geneva: Council on Health Research for Development; 2000.

27. Croxson B, Hanney S, Buxton M. Routine monitoring of performance: what makes health research and development different? *J Health Serv Res Policy* 2001;**6**:226–32.
28. Dixon S, Coleman P, Nicholl J, Brennan A, Touch S. Evaluation of the impact of a technology appraisal process in England: the South and West Development and Evaluation Committee. *J Health Serv Res Policy* 2003;**8**:18–24.
29. Drummond MF, Davies LM, Ferris FL III. Assessing the costs and benefits of medical research: the diabetic retinopathy study. *Soc Sci Med* 1992;**34**:973–81.
30. Eisenberg JM. Putting research to work: reporting and enhancing the impact of health services research. *Health Serv Res* 2001;**36** (2):x–xvii.
31. Elliott H, Popay J. How are policy makers using evidence? Models of research utilisation and local NHS policy making. *J Epidemiol Commun Health* 2000;**54**:461–8.
32. Ferguson B, Kelly P, Georgiou A, Barnes G, Sutherland B, Woodbridge B. Assessing payback from NHS reactive research programmes. *J Manage Med* 2000;**14**:25–36.
33. Hailey DM, Cowley DE, Dankiw W. The impact of health technology assessment. *Commun Health Stud* 1990;**14**:223–34.
34. Hailey DM, Crowe BL. The influence of health technology assessment on the diffusion of MRI in Australia. *Int J Technol Assess Health Care* 1993;**9**:522–9.
35. Hailey D. The influence of technology assessments by advisory bodies on health policy and practice. *Health Policy* 1993;**25**:243–53.
36. Hailey D, Corabian P, Harstall C, Schneider W. The use and impact of rapid health technology assessments. *Int J Technol Assess Health Care* 2000;**16**:651–6.
37. Hailey D. Elements of effectiveness for health technology assessment programs. *HTA initiative 9*. Edmonton: AHFMR; 2003.
38. Hall W. The contribution of research to Australian policy responses to heroin dependence 1990–2001: a personal retrospection. *Addiction* 2004;**99**:560–9.
39. Jacob R, Battista R. Assessing technology assessment. *Int J Technol Assess Health Care* 1993;**9**:564–72.
40. Jacob R, McGregor M. Assessing the impact of health technology assessment. *Int J Technol Assess Health Care* 1997;**13**:68–80.
41. Jonsson E, Banta HD, Schersten T. Health technology assessment and screening in Sweden. *Int J Technol Assess Health Care* 2001;**17**:380–8.
42. Kerridge L, Milne R, Robertson J, Gabbay J, Pattison K. Assessing the impact of 10 years of the NHS R&D Health Technology Assessment programme [abstract]. In *Proceedings of the First Annual Meeting of the Health Technology Assessment International (HTAi)*, Krakow, Poland, 30 May–2 June 2004. p. 31.
43. Lavis J, Ross S, McLeod C, Gildiner A. Measuring the impact of health research. *J Health Serv Res Policy* 2003;**8**:165–70.
44. McGregor M, Brophy M. End-user involvement in health technology assessment (HTA) development: a way to increase impact. *Int J Technol Assess Health Care*; 2005;**21**:263–7.
45. Molas-Gallart J, Tang P, Morrow S. Assessing the non-academic impact of grant-funded socio-economic research: results from a pilot study. *Res Eval* 2000;**9**:171–82.
46. National Institutes of Health. *Cost savings resulting from NIH research support*. 2nd. ed. Bethesda, MD: United States Department of Health and Human Services, National Institutes of Health; 1993.
47. NHS Executive South West. *An evaluation of the regional R&D project grants scheme*. Bristol: NHS Executive South West; 1999.
48. NHS Executive Trent. *An evaluation of regional R&D funding schemes in Trent*. Sheffield: NHS Executive Trent; 1997.
49. Raiten D, Berman S. *Can the impact of basic biomedical research be measured? A case study approach*. Bethesda, MD: Life Sciences Research Office, Federation of American Societies for Experimental Biology; 1993.
50. Royal Netherlands Academy of Arts and Sciences. *The societal impact of applied research towards a quality assessment system*. Amsterdam: Royal Netherlands Academy of Arts and Sciences; 2002.
51. Shah S, Ward JE. Outcomes from NHMRC public health research project grants awarded in 1993. *Aust N Z J Public Health* 2001;**25**:556–60.
52. Shani S, Siebzeiner M, I, Luxenburg O, Shemer J. Setting priorities for the adoption of health technologies on a national level: the Israeli experience. *Health Policy* 2000;**54**:169–85.
53. Spaapen J, Sylvain C. *Societal quality of research – towards a method for the assessment of the potential value of research for society*. SPSG Concept Paper 13. London: Science Policy Support Group; 1994.
54. Stryer D, Tunis S, Hubbard H, Clancy C. The outcomes of outcomes and effectiveness research: impacts and lessons from the first decade. *Health Serv Res* 2000;**35** (5 Pt 1):977–93.
55. Weisbrod BA. Costs and benefits of medical research: a case study of poliomyelitis. *J Political Econ* 1971;**79**:527–44.
56. Wisely J. *Assessment of the benefits from the National R&D Programme on primary/secondary care interface*. London: NHS London Regional Office; 2001.

57. Wisely J. National R&D Programme in the area of mother and child: programme report. London: NHS London Regional Office; 2001.
58. Yesudian CAK. Policy research in India: the case of regulating private health providers. Geneva: Alliance for Health Policy and Systems Research; 2001.
59. Hailey D, Juzwishin D. Managing external risks to health technology assessment programs. *Int J Technol Assess Health Care* 2006;**22**:429–35.
60. Weiss C.H. *Social science research and decision making*. Newark, NJ: Columbia University Press; 1977.
61. Weiss C. The many meanings of research utilisation. *Public Adm Rev* 1979;**39**:426–31.
62. Kogan M, Henkel M. *Government and research: the Rothschild experiment in a government department*. London: Heinemann; 1983.
63. Mushkin S. *Biomedical research: costs and benefits*. Cambridge, MA: Ballinger; 1979.
64. Molas-Gallart J, Tang P, Sinclair T, Morrow S, Martin B. *Assessing research impact on non-academic audiences. Report to the Economic and Social Research Council*. Brighton: SPRU, University of Sussex; 1999.
65. Johnston SC, Rootenberg JD, Katrak S, Smith WS, Elkins JS. Effects of a US National Institutes of Health programme of clinical trials on public health and costs. *Lancet* 2006;**367**:27.
66. Kingwell BA, Anderson GP, Duckett S, Hode EA, Jackson-Pulver LR, Khachigian LM. Evaluation of NHMRC funded research completed in 1992, 1997 and 2003. *Med J Aust* 2006;**184**:282–6.
67. Anderson R. The policy impact of population health surveys: an illustration of the measurement challenges using the NSW Health Survey. *Evid Policy* 2006;**2**:167–83.
68. Oliver TR, Singer RF. Health services research as a source of legislative analysis and input: the role of the California Health Benefits Program. *Health Serv Res* 2006;**41** (3, Part 11):1124–58.
69. Kuruvilla S, Mays N, Pleasant A, Walt G. Describing the impacts of health research: A Research Impact Framework. *BMC Health Services Research* 2006;**6**:134.
70. Kuruvilla S, Mays N, Walt G. Describing the impacts of health services and policy research. *J Health Serv Res Policy* 2007;**12**(Suppl 1):23–31.
71. Oortwijn WJ, Ligtoet A, Grant S, Wooding S, Hoorens S, Bouter LM, *et al*. *The returns from health care efficiency research in The Netherlands*. Santa Monica, CA: RAND; (in press).
72. Kwan P, Johnston J, Fung AYK, Chong DSY, Collins RA, Lo SV. A systematic evaluation of payback of publicly funded health and health services research in Hong Kong. *BMC Health Services Research* 2007;**7**:121.
73. Lavis J, Ross S, Hurley J, Hohendal J, Stoddart G, Woodward C, *et al*. Examining the role of health services research in public policymaking. *Milbank Q* 2005;**80**:125–54.
74. Buxton M, Hanney S, Packwood T, Roberts S, Youll P. Assessing benefits from Department of Health and National Health Service research and development. *Public Money Manage* 2000;**20**(4):29–34.
75. Wooding S, Hanney S, Buxton M, Grant J. Payback arising from research funding: evaluation of the Arthritis Research Campaign. *Rheumatology* 2005;**44**:1145–56.
76. Henshall C, Koch P, von Below G, Boer A, Conde-Oasagasti J, Dillon A, *et al*. Health technology assessment in policy and practice. Working Group 6 Report. *Int J Technol Assess Health Care* 2002;**18**:447–55.
77. Barbieri M, Drummond M. *The use of HTA in decision making: report to ECHTA/ECAHI Working Group*. York: University of York; 2002.
78. Cozzens S. In Anderson J, Fears R, editors. *Valuing and evaluating: assessment of the value of R&D in creating national and corporate prosperity*. Report on the Second Smith Kline Beecham Symposium. Oxford: Cross and Associates; 1996.
79. Black C. Evidence based policy: proceed with care. *BMJ* 2001;**323**:275–9.
80. Davies H, Nutley S, Walter I. *Approaches to assessing the non-academic impact of social science research*. Report of the ESRC symposium on assessing the non-academic impact of research, 12–13 May 2005. Available at: [www.esrcsocietytoday.ac.uk/ESRCInfoCentre/Forums/attach.aspx?a=64](http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/Forums/attach.aspx?a=64)
81. Hanney S, Home P, Frame I, Grant J, Green P, Buxton M. Identifying the impact of diabetes research. *Diabet Med* 2006;**23**:176–84.
82. Comroe JH, Dripps RD. Scientific basis for the support of biomedical science. *Science* 1976;**192**:105–11.
83. Grant G, Cottrell R, Cluzeau F, Fawcett E. Evaluating “payback” on biomedical research from papers cited in clinical guidelines: applied bibliometric study. *BMJ* 2000;**320**:1107–11.
84. Department of Health. Best research for best health: summary of responses to the consultation on a new national health research strategy. London: Department of Health; 2005. Available at: [www.dh.gov.uk/en/Consultations/Responsestoconsultations/DH\\_4124392](http://www.dh.gov.uk/en/Consultations/Responsestoconsultations/DH_4124392).
85. Hanney S, Grant J, Wooding S, Buxton M. Proposed methods for reviewing the outcomes of health research: the impact of funding by the UK’s Arthritis Research Campaign. *Health Res Policy Syst* 2004;**2**(4).

86. Yin R, Moore G. Lessons on the utilisation of research from nine case experiences in the natural hazards field. *Knowledge in Society* 1988;**1**:25–44.
87. Hanney S, Mugford M, Grant J, Buxton M. Assessing the benefits of health research: lessons from research into the use of antenatal corticosteroids for the prevention of neonatal respiratory distress syndrome. *Soc Sci Med* 2005; **60**:937–47.
88. Abacus International. *NICE guidance implementation tracking: data, sources, methodology & results*. Abacus International; 2005. Available at: [www.nice.org.uk/page.aspx?0=209966](http://www.nice.org.uk/page.aspx?0=209966)
89. Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I, *et al*. What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patient's notes and interviews. *BMJ* 2004;**329**:999–1006.
90. Kogan M, Henkel M., Hanney S. *Government and research: thirty years of evolution*. 2nd ed. Dordrecht: Springer; 2006.
91. Trostle J, Bronfman M, Langer A. How do researchers influence decision-makers? Case studies of Mexican policies. *Health Policy Planning* 1999;**14**:103–14.
92. Kalra L, Evans A, Perez I, Knapp M, Swift C, Donaldson N. A randomised controlled comparison of alternative strategies in stroke care. *Health Technol Assess* 2005;**9**(18).
93. Kalra L, Evans A, Perez I, Knapp M, Donaldson N, Swift C. Alternative strategies for stroke care: a prospective randomised controlled trial. *Lancet* 2000;**356**:894–9.
94. Duncan PW, Zorowitz R, Bates B, Choi JY, Glasberg MA, Glenn D, *et al*. Management of adult rehabilitation care: a clinical practice guideline. *Stroke* 2005;**36**:e100–43.
95. King M, Sibbald B, Ward E., Bower P, Lloyd M, Gabbay M, *et al*. Randomised controlled trial of nondirective counselling, cognitive-behaviour therapy and usual general practitioner care in the management of depression as well as mixed anxiety and depression in primary care. *Health Technology Assess* 2000;**4**(19).
96. Ward E, King M, Lloyd M, Bower P, Sibbald B, Farrelly S, *et al*. Randomised controlled trial of nondirective counselling, cognitive-behaviour therapy and usual general practitioner care for patients with depression. I: Clinical effectiveness. *BMJ* 2000;**321**:1383–8.
97. Department of Health. *National service framework for older people*. London: Department of Health; 2001.
98. Delaney B, Moayyedi P, Deeks J, Innes M, Soo S, Barton P, *et al*. The management of dyspepsia: a systematic review. *Health Technology Assess* 2000; **4**(39).
99. Moayyedi P, Soo S, Deeks J, Foreman D, Mason J, Innes M, *et al*. Systematic review and economic evaluation of *Helicobacter pylori* treatment for non-ulcer dyspepsia. *BMJ* 2000;**321**:659–64.
100. National Institute for Clinical Excellence. *Dyspepsia: managing dyspepsia in adults in primary care*. Evidence-Based Clinical Practice Guideline. London: NICE; 2004.
101. Hanney S, Davies A, Buxton M. Assessing benefits from health research projects: can we use questionnaire instead of case studies. *Res Eval* 1999;**8**:189–99.
102. Weiss C, Bucuvalas M. Truth tests and utility tests: decision-maker's frames of reference for social science research. *Am Soc Rev* 1980;**45**:302–13.
103. Wanke M, Juzwihin D, Thornley R, Chan L. *An exploratory review of evaluations of health technology assessment agencies*. Edmonton: AHFMR; 2006.



# Appendix I

## Key references identified from literature search

- Adamson TA. The productivity and impact of psychiatric research on development of mental health services in Nigeria. *W Afr J Med* 1998;**17**:243–7.
- Afseth J. On the value of research in health care. *Eur J Oral Sci* 1997;**105**:454–6.
- Alberta Heritage Foundation for Medical Research (AHFMR). *A study of the impact of 2000–2001 HTA products*. Edmonton: AHFMR; 2002.
- Alberta Heritage Foundation for Medical Research (AHFMR). *A study of the impact of 2001–2002 health technology assessment products: final report*. Edmonton: AHFMR; 2003.
- Alberta Heritage Foundation for Medical Research (AHFMR). *Assessment of health research fund outputs and outcomes: 1995–2003*. Edmonton: AHFMR; 2003.
- Anderson MCJ, Swan BMHBM. The use of research in local health services agencies. *Soc Sci Med* 1999;**49**:1007–19.
- Aries NR, Sclar ED. The economic impact of biomedical research: a case study of voluntary institutions in the New York metropolitan region. *J Health Politics Policy Law* 1998;**23**:175–93.
- Australian Society for Medical Research. *Australian case studies, series I & II*. Sydney: Australian Society for Medical Research; 2003. URL: <http://www.asmr.org.au/news/>
- Baker M, Kirk S, Donaldson L. *Research and development for the NHS: evidence, evaluation and effectiveness*. Abingdon: Radcliffe Medical Press; 2001.
- Banta D, Oortwijn WJ. Conclusions: health technology assessment and health care in the European Union. *Int J Technol Assess Health Care* 2000;**16**:626–35.
- Banta HD, Oortwijn WJ, Beekun van WT. *The organization of health care technology assessment in The Netherlands*. The Hague: Rathenau Instituut; 1995.
- Banta D, Oortwijn WJ. Health technology assessment and screening in The Netherlands: case studies of mammography in breast cancer, PSA screening in prostate cancer, and ultrasound in normal pregnancy. *Int J Technol Assess Health Care* 2001;**17**:369–79.
- Banta HD. Health Policy, health technology assessment and screening in Europe. *Int J Technol Assess Health Care* 2001;**17**:409–17.
- Barbieri M, Drummond M. *The use of HTA in decision making: Report to ECHTA/ECAHI Working Group*. York: University of York; 2002.
- Battista R. Health care technology assessment: linking science and policy-making. *Can Med Assoc J* 1992;**146**:461–2.
- Battista RN, Banta HD, Jonsson E, Hodge M, Gelband H. Lessons from the eight countries. *Health Policy* 1994;**30**:397–421.
- Berg M, van der Grinten T, Klazinga N. Technology assessment, priority setting, and appropriate care in Dutch health care [published erratum appears in *Int J Technol Assess Health Care* 2004;**20**:564]. *Int J Technol Assess Health Care* 2004;**20**:35–43.
- Berridge V, Stanton J. Science and policy: historical insights. *Soc Sci Med* 1999;**9**:1133–8.
- Bircumshaw D. The utilisation of research findings in clinical nursing practice. *J Adv Nurs* 1990;**15**:1272–80.
- Bjerregaard P, Curtis T. Impact of public health research in Greenland. *Int J Circumpolar Health* 2004;**63**:214–20.
- Boer A. Assessment and regulation of health care technology: the Dutch experience. *Int J Technol Assess Health Care* 1999;**15**:638–48.
- Bogenscheider K, Oleson JR, Linney KD, Mills J. Connecting research and policymaking: implications for theory and practice from the family impact seminars. *Fam Relations* 2000;**48**:327–39.
- Bonsel GJ. Orthoptic liver transplant in The Netherlands. The results and impact of a medical technology assessment. *Health Policy* 1990;**16**:147–61.
- Borras JM, Granados A, Escarrabill J, de-Lissovoy G. Complex decisions about an uncomplicated therapy: reimbursement for long-term oxygen therapy in Catalonia (Spain). *Health Policy* 1996;**35**:53–9.
- Bos M, Carlsson P, Kooj S, Liaropoulos L, Sampietro-Colom L, Schilling J, et al. *Technology assessment and coverage policy: the case of invasive cardiology therapy in five European countries*. Barcelona: Catalan Agency for Health Technology Assessment and Research; 1997.
- Bradley EH, Webster TR, Baker D, Schlesinger M, Inouye SK, Barth MC, et al. Translating research into practice: speeding the adoption of innovative health care programs. *Issue Brief (Commonwealth Fund)* 2004;**724**:1–12.
- Britton M, Jonsson E. Impact of health technology assessments. Some experiences of SBU. *Int J Technol Assess Health Care* 2002;**18**:824–31.
- Brorsson B, Arvidsson S. The effect of dissemination of recommendations on use. Preoperative routines in Sweden, 1989–91. *Int J Technol Assess Health Care* 1997;**13**:547–52.
- Brownson RC, Simoes EJ. Measuring the impact of prevention research on public health practice. *Am J Prev Med* 1999;**16** (3 Suppl):72–9.

- Bucci VA. Health outcomes research: its influence on clinical decision making and the development of new imaging technologies. *AJNR Am J Neuroradiol* 1991; **12**:397–9.
- Burns A, Charlwood P, Darling H, Fox D, Greenfield L, Hamlyn L, *et al.* Better information? Better outcomes? *The use of health technology assessment and clinical effectiveness data in healthcare purchasing decisions in the United Kingdom and the United States*. New York: Milbank Memorial Fund; 2000. pp. 1–26.
- Buxton M, Hanney S, Packwood T, Roberts S, Youll P. Assessing benefits from Department of Health and National Health Service research and development. *Public Money Manage* 2000; **20**:29–34.
- Buxton M, Schneider W. Assessing the 'payback' from AHFMR-funded research. Edmonton: AHFMR; 1999.
- Buxton M, Hanney S. Evaluating the NHS Research and Development Programme: will the programme give value for money? *Journal R Soc Med* 1998; **91** (Suppl 35):2–6.
- Buxton M, Hanney S, Jones T. Estimating the economic value to societies of the impact of health research: a critical review. *Bull World Health Organ* 2004; **82**:733–9.
- Buxton M, Hanney S. How can payback from health services research be assessed?. *J Health Serv Res Policy* 1996; **1**:35–43.
- Caplan N, Morrison A, Stambaugh RJ. *The use of social science knowledge in policy decisions at the national level: a report to respondents*. Ann Arbor, MI: University of Michigan; 1975.
- Carlsson P. Health technology assessment and priority setting for health policy in Sweden. *Int J Technol Assess Health Care* 2004; **20**:44–54.
- Cavanagh SJ, Tross G. Utilizing research findings in nursing: policy and practice considerations. *J Adv Nurs* 1996; **24**:1083–8.
- Cochrane Collaboration/Milbank Memorial Fund. *Informing judgement: case studies of health policy and research in six countries*. New York: Cochrane Collaboration/Milbank Memorial Fund; 2001.
- COHRED Working Group on Research to Action and Policy. *Lessons in research to action and policy – case studies from seven countries*. Geneva: Council on Health Research for Development; 2000.
- Coleman P, Nicholl J. Influence of evidence-based guidance on health policy and clinical practice in England. *Qual Health Care* 2001; **10**:229–37.
- Comroe JH, Dripps RD. Scientific basis for the support of biomedical science. *Science* 1976; **192**:105–11.
- Council of Health Research and Development (COHRED) Working Group on Research to Action and Policy. *The role of research in child health policy and programs in Pakistan*. Geneva: Council of Health Research and Development (COHRED); 2000.
- Cranovsky R, Matillon Y, Banta D. EUR-ASSESS project subgroup on coverage. *Int J Technol Assess Health Care* 1997; **13**:287–332.
- Cronin E, Normand C. *Methodologies for evaluating R&D programmes – a literature review*. London: LSHTM; 1999.
- Croxson B, Hanney S, Buxton M. Routine monitoring of performance: what makes health research and development different? *J Health Serv Res Policy* 2001; **6**:226–32.
- Dash P, Gowman N, Traynor M. Increasing the impact of health services research. *BMJ* 2003; **327**:1339–41.
- Davis P, Howden-Chapman P. Translating research findings into health policy. *Soc Sci Med* 1996; **43**:865–72.
- Degeling P, Maxwell SA, Kennedy J, Coyle B, Macbeth F. The impact of CHI: evidence from Wales. *Qual Primary Care* 2003; **11**:147–54.
- Department of Health. *Review of the research and development taking place in the London Postgraduate Special Health Authorities*. A Report by the Review Advisory Committee. London: HMSO; 1993.
- Department of Health. *Report of the NHS Technology Assessment Programme 1996*. London: Department of Health; 1996.
- Dixon S, Coleman P, Nicholl J, Brennan A, Touch S. Evaluation of the impact of a technology appraisal process in England: the South and West Development and Evaluation Committee. *J Health Serv Res Policy* 2003; **8**:18–24.
- Dobbins M, Ciliska D, Cockerill R, Barnsley J, DiCenso A. A framework for the dissemination and utilization of research for health-care policy and practice. *Online J Knowledge Synth Nurs* 2002; **9**:7.
- Drummond M, Cooke J, Walley T. Economic evaluation under managed competition: evidence from the UK. *Soc Sci Med* 1997; **49**:583–95.
- Drummond M, Brown R, Fendrick AM, Fullerton P, Neumann P, Taylor R, *et al.* Use of pharmacoeconomic information – report of the ISPOR Task Force on the use of pharmacoeconomic/health economic information in healthcare decision making. *Value Health* 2003; **4**:407–16.
- Drummond M, Weatherly H. Implementing the findings of health technology assessments. If the CAT got out of the bag, can the TAIL wag the dog? *Int J Technol Assess Health Care* 2000; **16**:1–12.
- Drummond MF, Davies LM, Ferris FL III. Assessing the costs and benefits of medical research: the diabetic retinopathy study. *Soc Sci Med* 1992; **34**:973–81.
- Durrheim DN, Speare R, Harries AD. Research that influences policy and practice – characteristics of operational research to improve malaria control in Mpumalanga Province, South Africa. *Malaria J* 2002; **1**(9).
- Duthie T, Trueman P, Chancellor J, Diez L. Research into the use of health economics in decision making in the United Kingdom, phase II: is health economics “for good or evil”? *Health Policy* 1999; **46**:143–57.

- Eagar K, Cromwell D, Owen A, Senior K, Gordon R, Green J. Health Services research and development in practice: an Australian example. *J Health Serv Res Policy* 2003;**8** (Suppl 2):7–13.
- Eisenberg JM. Putting research to work: reporting and enhancing the impact of health services research. *Health Serv Res* 2001;**36** (2):x–xvii.
- Elliott H, Popay J. How are policy makers using evidence? Models of research utilisation and local NHS policy making. *J Epidemiol Commun Health* 2000;**54**:461–8.
- Elsinga E, Rutten FFH. Economic evaluation in support of national health policy: the case of The Netherlands. *Soc Sci Med* 1997;**45**:605–20.
- Faisst K, Schilling J, Koch P. Health technology assessment of three screening methods in Switzerland. *Int J Technol Assess Health Care* 2001;**17**:389–99.
- Farmer J, Chesson R. Nice lessons to be learned. *Int J Technol Assess Health Care* 2001;**17**:222–35.
- Favaretti C, De Pieri P. Mammography, routine ultrasonography in pregnancy, and PSA screenings in Italy. *Int J Technol Assess Health Care* 2001;**17**:358–68.
- Ferguson B, Kelly P, Georgiou A, Barnes G, Sutherland B, Woodbridge B. Assessing payback from NHS reactive research programmes. *J Manage Med* 2000;**14**:25–36.
- Ferguson JH. NIH consensus conferences: dissemination and impact. *Ann N Y Acad Sci* 1993;**31**:703:180–98.
- Ferlie E, Fitzgerald L, Wood M. Getting evidence into clinical practice: an organisational behaviour perspective. *J Health Serv Res Pol* 2000;**5**:96–102.
- Fletcher A. The NHS research and development programme and its impact on health care library and information resource centres. *Health Libraries Rev* 1995;**12**:243–7.
- Florin D. Scientific uncertainty and the role of expert advice: the case of health checks for coronary heart disease prevention by general practitioners in the UK. *Soc Sci Med* 1999;**49**:1269–83.
- Freemantle N, Mason JM, Watt I. Evidence into practice. Prescribing selective serotonin reuptake inhibitors. *Int J Technol Assess Health Care* 1998;**14**:387–91.
- Fudenberg H, editor. *Biomedical institutions, biomedical funding and public policy*. New York: Plenum Press; 1983.
- Gelijns AC, Rosenberg N, Moskowitz AJ. Capturing the unexpected benefits of medical research. *N Engl J Med* 1998;**339**:693–8.
- Gerhardus A, Dujardin M, Kiet P, Siddiqi S, Sauerborn R. *A methodology to assess the use of research for health policy development*. Discussion Paper 04. Heidelberg: University of Heidelberg Medical School; 2000.
- Gerhardus A, Kilemann K, Sanou A. Lessons in research to action and policy: case studies from seven different countries. Geneva: Council of Health Research and Development (COHRED) Working Group on Research to Action and Policy; 2000.
- Gerhardus A, Dintsios C. The impact of HTA reports on health policy – a systematic review. Cologne: German Agency of Health Technology Assessment at German Institute for Medical Documentation and Information (DAHTA) (DIMDI) 2004 (in German).
- Gerhardus ADCD. A systematic review of the impact of HTA-reports on health policy shows lack of systematic evaluation and evaluation tools. In *The Challenge of Collaboration. 18th Annual Meeting of the International Society of Technology Assessment in Health Care*, 9–12 June, 2002, Berlin; 2002; p. 76.
- Ginzberg E. *Health services research: key to health policy*. Cambridge, MA: Harvard University Press; 1991.
- Gold RS, Atkinson NL. Importance of health education research to health education. *Am J Health Behav* 2001;**25**:301–10.
- Granados A, Jonsson E, Banta HD, Bero L, Bonair A, Cochet C, et al. EUR-ASSESS Project Subgroup Report on Dissemination and Impact. *Int J Technol Assess Health Care* 1997;**13**:220–86.
- Grant J. Evaluating the outcomes of biomedical research on health care. *Res Eval* 1999;**8**:33–8.
- Grant J, Allen L. Evaluating high risk research: an assessment of the Wellcome Trust's Sir Henry Wellcome Commemorative Awards for Innovative Research. *Res Eval* 1999;**8**:201–4.
- Grant J, Cottrell R, Cluzeau F, Fawcett G. Evaluating “payback” on biomedical research from papers cited in clinical guidelines: applied bibliometric study. *BMJ* 2000;**320**:1107–11.
- Gray JA. Evidence-based screening in the United Kingdom. *Int J Technol Assess Health Care* 2001;**17**:400–8.
- Hailey D. Australian economic evaluation and government decisions on pharmaceuticals, compared to assessment of other health technologies. *Soc Sci Med* 1997;**45**:563–81.
- Hailey D. The influence of technology assessments by advisory bodies on health policy and practice. *Health Policy* 1993;**25**:243–53.
- Hailey D, Corabian P, Harstall C, Schneider W. The use and impact of rapid health technology assessments. *Int J Technol Assess Health Care* 2000;**16**:651–6.
- Hailey D. Providing information on emerging health technologies to provincial decision makers: a pilot project. *Health Policy* 2001;**58**:15–26.
- Hailey D. *Elements of effectiveness for health technology assessment programs*. HTA Initiative 9. Edmonton: AHFMR; 2003.
- Hailey DM, Cowley DE, Dankiw W. The impact of health technology assessment. *Commun Health Stud* 1990;**14**:223–34.
- Hailey DM, Crowe BL. The influence of health technology assessment on the diffusion of MRI in Australia. *Int J Technol Assess Health Care* 1993;**9**:522–9.

Hall W. The contribution of research to Australian policy responses to heroin dependence 1990–2001: a personal retrospection. *Addiction* 2004;**99**:560–9.

Hannan A, Enright H, *et al.* *Using research the results of a pilot study comparing teachers, general practitioners and surgeons.* 1996. URL: <http://www.leeds.ac.uk/educol/documents/000000851.htm>

Hanney S, Mugford M, Grant J, Buxton M. Assessing the benefits of health research: lessons from research into the use of antenatal corticosteroids for the prevention of neonatal respiratory distress syndrome. *Soc Sci Med* 2005;**60**:937–47.

Hanney S, Buxton M, Brown J, Bryan S, Ratcliffe J, Walden L, *et al.* *Assessing payback from Department of Health research and development: second report: Vol 2: ten further case studies.* HERG Research Report 24. Uxbridge: Brunel University; 1997.

Hanney S, Davies A, Buxton M. Evaluating the benefits from health research projects. *Res Eval* 1999;**8**:189–99.

Hanney S, Soper B, Buxton M. *Evaluation of the NHS R&D Methods Programme.* HERG Research Report 29. Uxbridge: Brunel University; 2003.

Hanney SR, Gonzalez-Block MA, Buxton MJ, Kogan M. The utilisation of health research in policy-making: concepts, examples and methods of assessment. *Health Res Policy Syst* 2003;**1**(2).

Hanney S, Packwood T, Buxton M. Evaluating the benefits from health research and development centres: a categorization, a model and examples of application. *Evaluation* 2000;**6**:137–60.

Health Canada. *Report on the evaluation of the National Health Research and Development Program.* Ottawa: Program Evaluation Division, Program Audit & Review Directorate; 1994.

Henshall C, Koch P, von Below G, Boer A, Conde-Oasagasti J, Dillon A, *et al.* Health technology assessment in policy and practice. Working Group 6 Report. *Int J Technol Assess Health Care* 2002;**18**:447–55.

Hivon M, Lehoux P, Denis J-L, Tailliez S. Use of health technology assessment in decision making: Co-responsibility of users and producers? *Int J Technol Assess Health Care* 2005;**21**:268–75.

Hoffmann C, Stoykova BA, Nixon J, Glanville JM, Misso K, Drummond MF. Do health-care decision makers find economic evaluations useful? The findings of focus group research in UK health authorities. *Value Health* 2002;**5**:71–8.

Hutton J, Backhouse R. *Payback from research and development: final report.* London: Battelle Medical Technology and Policy Research Centre; 1994.

Ibbotson SL, Long AF, Sheldon TA, Mason J. An initial evaluation of effective health care bulletins as instruments of effective dissemination. *J Manage Med* 1993;**7**:48–57.

Ijzerman M, Reuzel R, Severens H. Pre-assessment to assess the match between cost-effectiveness results and

decision makers' information needs: an illustration using two cases in rehabilitation medicine in The Netherlands. *Int J Technol Assess Health Care* 2003;**19**:17–27.

Innvaer S, Vist G, Trommald M, Oxman A. Health policy-makers' perceptions of their use of evidence. *J Health Serv Res Policy* 2002;**7**:239–44.

Jacob R, McGregor M. Assessing the impact of health technology assessment. *Int J Technol Assess Health Care* 1997;**13**:68–80.

Johri M, Lehoux P. The great escape? Prospects for regulating access to technology through health technology assessment. *Int J Technol Assess Health Care* 2003;**19**:179–93.

Jonsson E, Banta HD, Schersten T. Health technology assessment and screening in Sweden. *Int J Technol Assess Health Care* 2001;**17**:380–8.

Kerridge L, Milne R, Robertson J, Gabbay J, Pattison K. Assessing the impact of 10 years of the NHS R&D Health Technology Assessment programme. In *Proceedings of the First Annual Meeting of the Health Technology Assessment International (HTAi)*, Krakow, Poland, 30 May–2 June 2004, p. 31.

Kingdon T. The relevance of research to policy formulation: an Australian perspective. *Addiction* 1993;**88** Suppl:61S–65S.

Kostoff RN. Research impact quantification. *R&D Manage* 1994;**24**:206–18.

Kothari A, Birch S, Charles C. "Interaction" and research utilisation in health policies and programs: does it work? *Health Policy* 2005;**71**:117–25.

Lavis J, Ross S, Hurley J, Hohendal J, Stoddart G, Woodward C, *et al.* Examining the role of health services research in public policymaking. *Milbank Q* 2002;**80**:125–54.

Lavis J, Ross S, McLeod C, Gildiner A. Measuring the impact of health research. *J Health Serv Res Policy* 2003;**8**:165–70.

Lavis JN, Ross SE. *The role of health services research in Canadian provincial policy-making.* CHSRT Report 1997–021. Hamilton: McMaster University; 2000.

Lomas J. Finding audiences, changing beliefs: the structure of research use in Canadian health policy. *J Health Politics Policy Law* 1990;**15**:525–42.

Lomas J. Using linkage and exchange to move research into policy at a Canadian Foundation. *J Health Politics Policy Law* 2000;**19**:236–40.

Lomas J. Connecting research and policy. *ISUMA Can J Policy Res* 2000;**1**:140–4.

Luce B, Brown RE. The use of technology assessment by hospitals, health technology maintenance organisations, and third-party payers in the United States. *Int J Technol Assess Health Care* 1995;**11**:79–92.

Lyles A, Luce B, Rentz A. Managed care pharmacy, socio-economic assessments and drug adoption decisions. *Soc Sci Med* 1997;**45**:511–21.

- Macintyre S, Chalmers J, Horton R, Smith R. Using evidence to inform health policy: case study. *BMJ* 2001; **322**:222–5.
- McDaid D, Cookson R, ASTEC Group. Evaluating health care interventions in the European Union. *Health Policy* 2003; **63**:133–9.
- Mitton C. Health technology assessment and policy decisions on hyperbaric oxygen treatment. *Int J Technol Assess Health Care* 1999; **15**:661–70.
- Molas-Gallart J, Tang P, Morrow S. Assessing the non-academic impact of grant-funded socio-economic research: results from a pilot study. *Res Eval* 2000; **9**:171–82.
- Moodley J, Jacobs M. *Research to action and policy: combating vitamin A deficiencies in South Africa*. Geneva: Council of Health Research and Development (COHRED) Working Group on Research to Action and Policy; 2000.
- Mousiama TI. Health technology assessment in the area of prevention – selected screening cases in Greece. *Int J Technol Assess Health Care* 2001; **17**:338–57.
- Murphy K, Topel R. *Measuring the gains from medical research. An economic approach*. Chicago: University of Chicago Press; 2003.
- Mushkin S. *Biomedical research: costs and benefits*. Cambridge, MA: Ballinger; 1979.
- NHS Executive South West. *An evaluation of the regional R&D project grants scheme*. Bristol: NHS Executive South West; 1999.
- National Institutes of Health. *Cost savings resulting from NIH research support*. 2nd ed. Bethesda, MD: United States Department of Health and Human Services, National Institutes of Health; 1993.
- National Institutes of Health. *FY 2000 research program outcomes assessment material*. 2nd ed. Bethesda, MD: United States Department of Health and Human Services, National Institutes of Health; 2000.
- Nutley S, Percy-Smith J, Solebury W. *Models of research impact: a cross sector review of literature and practice*. London: Learning and Skills Research Centre; 2003.
- Oh CH. Explaining the impact of policy information on policy making. *Knowledge and Policy: International Journal of Knowledge Transfer and Utilisation* 1997; **10**:25–55.
- Oliver A, Mossialos E, Robinson R. Health technology assessment and its influence on health-care priority setting. *Int J Technol Assess Health Care* 2004; **20**:1–10.
- Orosz E. The impact of social science research on health policy. *Soc Sci Med* 1994; **39**:1287–93.
- Pardes H, Manton J, Lander E, Tolley H, Ullian A, Palmer H. Effects of medical research on health care and the economy. *Science* 1999; **283**:36–7.
- Peipert JF. The economic value of medical research: is it worth the investment? *Obstet Gynecol* 2002; **99** (5 Pt 1): 835–40.
- Perleth M, Busse R, Gibis B, Brand A. Evaluation of preventive technologies in Germany. *Int J Technol Assess Health Care* 2001; **17**:329–37.
- Pritchard C, George C, Raine J. *Capturing the unexpected benefits of medical research*. London: Office of Health Economics; 2001.
- Raiten D, Berman S. *Can the impact of basic biomedical research be measured? A case study approach*. Bethesda, MD: Life Sciences Research Office, Federation of American Societies for Experimental Biology; 1993.
- Ramsey SD, Sullivan SD. Evidence, economics, and emphysema: Medicare's long journey with lung volume reduction surgery. *Health Affairs* 2005; **24**:55–66.
- Richardson A, Jackson C, Sykes W. *Taking research seriously: means of improving and assessing the use and dissemination of research*. London: HMSO; 1990.
- Roos N, Shapiro E. From research to policy: what have we learned? *Med Care* 1999; **37**:JS291–305.
- Rosen R. Applying research to health care policy and practice: medical and managerial views on effectiveness and the role of research. *Journal Health Serv Res Policy* 2000; **5**:103–8.
- Rosenberg L. Exceptional economic returns on investments in medical research. *Med J Aust* 2002; **177**:368–71.
- Ross J. The use of economic evaluation in health care: Australian decision makers' perceptions. *Health Policy* 1995; **31**:103–10.
- Royal Netherlands Academy of Arts and Sciences. *The societal impact of applied research towards a quality assessment system*. Amsterdam: Royal Academy of Arts and Sciences; 2002.
- Sauerborn R, Nitayarumphong S, Gerhardus A. Strategies to enhance the use of health systems research for health sector reform. *Trop Med Int Health* 1999; **4**:827–35.
- Shah S, Ward JE. Outcomes from NHMRC public health research project grants awarded in 1993. *Aust N Z J Public Health* 2001; **25**:556–60.
- Shani S. Setting priorities for the adoption of health technologies on a national level – the Israeli experience. *Health Policy* 2000; **54**:169–85.
- Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I, et al. What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes and interviews. *BMJ* 2004; **329**: 999–1006.
- Silverstein S, Garrison H, Heinig S. A few basic economic facts about research in the medical and related life sciences. *FASEB J* 1995; **9**:833–40.
- Sloan F, Grabowski H. Introduction and overview. *Soc Sci Med* 1997; **45**:505–10.

- Smith R. Measuring the social impact of research: difficult but necessary. [Editorial]. *BMJ* 2001;**323**:528.
- Spaapen J, Sylvain C. *Societal quality of research – towards a method for the assessment of the potential value of research for society*. SPSP Concept Paper 13. London: Science Policy Support Group; 1994.
- Stemerding DVB. Maternal serum screening, political decision-making, and social learning. *Health Policy* 2001; **56**:111–25.
- Stevens A, Milne R. Health technology assessment in England and Wales. *Int J Technol Assess Health Care* 2004; **20**:11–24.
- Stocking B. Why research findings are not used by commissions – and what can be done about it. *J Public Health Med* 1997;**17**:380–2.
- Stryer D, Tunis S, Hubbard H, Clancy C. The outcomes of outcomes and effectiveness research: impacts and lessons from the first decade. *Health Serv Res* 2000; **35** (5 Pt 1):977–93.
- Trostle J, Bronfman M, Langer A. How do researchers influence decision-makers? Case studies of Mexican policies. *Health Policy Planning* 1999;**14**:103–14.
- Tugwell P, Sitthi-Amorn C, O'Connor A, Hatcher-Roberts J, Bergevin Y, Wolfson M. Technology assessment: old, new, and needs-based. *Int J Technol Assess Health Care* 1995;**11**:650–61.
- van den Heuvel W, Weiring H R, van den Heuvel L. Utilisation of medical technology assessment in health policy. *Health Policy* 1997;**42**:211–21.
- van Rossum W. Decision-making and medical technology assessment: three Dutch cases. *Knowledge Policy* 1991;**4**:107–24.
- van Weel C. Biomedical science matters for people – so its impact should be better assessed. *Lancet* 2002; **360**:1034–5.
- Vermeulen V, Coppens K, Kesteloot K. Impact of health technology assessment on preventive screening in Belgium: case studies of mammography in breast cancer, PSA screening in prostate cancer, and ultrasound in normal pregnancy. *Int J Technol Assess Health Care* 2001;**17**:316–28.
- Vingilis E, Hartford K, Schrecker T, Mitchell B, Lent B, Bishop J, et al. Integrating knowledge generation with knowledge diffusion and utilization: a case study analysis of the Consortium for Applied Research and Evaluation in Mental Health. *Can J Public Health* 2003; **94**:468–71.
- von Lengerke T, Rutten A, Vinck J, Abel T, Kannas L, Luschen G, et al. Research utilization and the impact of health promotion policy. *Soz Präventivmed* 2004;**49**:185–97.
- Walt G. How far does research influence policy? *Eur J Public Health* 1994;**4**:233–5.
- Walter I, Davies H, Nutley S. Increasing research impact through partnerships: evidence from outside health care. *J Health Serv Res Policy* 2003;**8** Suppl 2:58–61.
- Wathen B, Dean T. An evaluation of the impact of NICE guidance on GP prescribing. *Br J Gen Pract* 2004; **54**:103–7.
- Weisbrod BA. Costs and benefits of medical research: a case study of poliomyelitis. *J Political Econ* 1971; **79**:527–44.
- Weisbrod BA. *Economics and medical research*. Washington DC: American Enterprise Institute for Public Policy research; 1983.
- Weiss C, Bucuvalas M. Truth tests and utility tests: decision-maker's frames of reference for social science research. *Am Soc Rev* 1980;**45**:302–13.
- Wild C. Screening in Austria: the cases of mammography, PSA testing and routine use of ultrasound in pregnancy. *Int J Technol Assess Health Care* 2001;**17**:305–15.
- Williamson P. From dissemination to use: management and organisational barriers to the application of health services research findings. *Health Bull* 1992;**50**:78–86.
- Wisely J. *Assessment of the benefits from the National R&D Programme on Primary/Secondary Care Interface*. Unpublished report. London: NHS London Regional Office; 2001.
- Wisely J. *National R&D programme in the area of mother and child: programme report*. London: NHS London Regional Office; 2001.
- Wooding S, Hanney S, Buxton M, Grant J. *The returns from arthritis research. Volume 1: approach, analysis and recommendations*. Cambridge: RAND, Europe; 2004.
- Yesudian C A K. *Policy research in India: the case of regulating private health providers*. Geneva: The Alliance for Health Policy and Systems Research; 2001.

## Appendix 2

### Literature review 'evidence base'

- Adamson TA. The productivity and impact of psychiatric research on development of mental health services in Nigeria. *W Afr J Med* 1998;**17**:243–7.
- Alberta Heritage Foundation for Medical Research (AHFMR). *A study of the impact of 2000–2001 HTA products*. Edmonton: AHFMR; 2002.
- Alberta Heritage Foundation for Medical Research (AHFMR). *Assessment of health research fund outputs and outcomes: 1995–2003*. Edmonton: AHFMR; 2003.
- Alberta Heritage Foundation for Medical Research (AHFMR). *A study of the impact of 2001–2002 health technology assessment products: final report*. Edmonton: AHFMR; 2003.
- Bjerregaard P, Curtis T. Impact of public health research in Greenland. *Int J Circumpolar Health* 2004;**63**:214–20.
- Bos M, Carlsson P, Kooj S, Liaropoulos L, Sampietro-Colom L, Schilling J, et al. *Technology assessment and coverage policy: the case of invasive cardiology therapy in five European countries*. Barcelona: Catalan Agency for Health Technology Assessment and Research; 1997.
- Britton M, Jonsson E. Impact of health technology assessments. Some experiences of SBU. *Int J Technol Assess Health Care* 2002;**18**:824–31.
- Brownson RC, Simoes EJ. Measuring the impact of prevention research on public health practice. *Am J Prev Med* 1999;**16** (3 Suppl):72–9.
- Buxton M, Hanney S. How can payback from health services research be assessed? *J Health Serv Res Policy* 1996;**1**:35–43.
- Buxton M, Hanney S, Packwood T, Roberts S, Youll P. *Assessing the benefits from North Thames research and development*. HERG Research Report 25. Uxbridge: Brunel University Health Economics Research Group; 1999.
- Buxton M, Schneider W. *Assessing the 'payback' from AHFMR-funded research*. Edmonton: AHFMR; 1999.
- Council of Health Research and Development (COHRED) Working Group on Research to Action and Policy. *The role of research in child health policy and programs in Pakistan*. Geneva: Council of Health Research and Development (COHRED); 2000.
- Croxson B, Hanney S, Buxton M. Routine monitoring of performance: what makes health research and development different? *J Health Serv Res Policy* 2001;**6**:226–32.
- Dixon S, Coleman P, Nicholl J, Brennan A, Touch S. Evaluation of the impact of a technology appraisal process in England: the South and West Development and Evaluation Committee. *J Health Serv Res Policy* 2003;**8**:18–24.
- Drummond MF, Davies LM, Ferris FL, III. Assessing the costs and benefits of medical research: the diabetic retinopathy study. *Soc Sci Med* 1992;**34**:973–81.
- Eisenberg JM. Putting research to work: reporting and enhancing the impact of health services research. *Health Serv Res* 2001;**36** (2):x–xvii.
- Elliott H, Popay J. How are policy makers using evidence? Models of research utilisation and local NHS policy making. *J Epidemiol Commun Health* 2000;**54**:461–8.
- Ferguson B, Kelly P, Georgiou A, Barnes G, Sutherland B, Woodbridge B. Assessing payback from NHS reactive research programmes. *J Manage Med* 2000;**14**:25–36.
- Hailey DM, Cowley DE, Dankiw W. The impact of health technology assessment. *Community Health Stud* 1990;**14**:223–34.
- Hailey D. The influence of technology assessments by advisory bodies on health policy and practice. *Health Policy* 1993;**25**:243–53.
- Hailey DM, Crowe BL. The influence of health technology assessment on the diffusion of MRI in Australia. *Int J Technol Assess Health Care* 1993;**9**:522–9.
- Hailey D, Corabian P, Harstall C, Schneider W. The use and impact of rapid health technology assessments. *Int J Technol Assess Health Care* 2000;**16**:651–6.
- Hailey D. *Elements of effectiveness for health technology assessment programs*. HTA Initiative No. 9. Edmonton: AHFMR; 2003.
- Hall W. The contribution of research to Australian policy responses to heroin dependence 1990–2001: a personal retrospection. *Addiction* 2004;**99**:560–9.
- Hanney S, Soper B, Buxton M. *Evaluation of the NHS R&D Methods Programme*. HERG Research Report 29. Uxbridge: Brunel University; 2003.
- Jacob R, Battista R. Assessing technology assessment. *Int J Technol Assess Health Care* 1993;**9**:564–72.
- Jacob R, McGregor M. Assessing the impact of health technology assessment. *Int J Technol Assess Health Care* 1997;**13**:68–80.
- Jonsson E, Banta HD, Schersten T. Health technology assessment and screening in Sweden. *Int J Technol Assess Health Care* 2001;**17**:380–8.
- Kerridge L, Milne R, Robertson J, Gabbay J, Pattison K. Assessing the impact of 10 years of the NHS R&D Health Technology Assessment programme [abstract]. In *Proceedings of the First Annual Meeting of the Health*

*Technology Assessment International (HTAi)*, Krakow, Poland, 30 May–2 June 2004; p. 31.

Lavis J, Ross S, McLeod C, Gildiner A. Measuring the impact of health research. *J Health Serv Res Policy* 2003; **8**:165–70.

McGregor M, Brophy J. End-user involvement in health technology assessment (HTA) development: a way to increase impact. *Int J Technol Assess Health Care* 2004; **21**:263–7.

Molas-Gallart J, Tang P, Morrow S. Assessing the non-academic impact of grant-funded socio-economic research: results from a pilot study. *Res Eval* 2000; **9**:171–82.

National Institutes of Health. *Cost savings resulting from NIH research support*. 2nd ed. Bethesda, MD: United States Department of Health and Human Services, National Institutes of Health; 1993.

NHS Executive South West. *An evaluation of the regional R&D project grants scheme*. Bristol: NHS Executive South West; 1999.

NHS Executive Trent Region. *An evaluation of regional R&D funding schemes in Trent*. Sheffield: NHSE Executive Trent; 1997.

Raiten D, Berman S. *Can the impact of basic biomedical research be measured? A case study approach*. Bethesda, MD: Life Sciences Research Office, Federation of American Societies for Experimental Biology; 1993.

Royal Netherlands Academy of Arts and Sciences. *The societal impact of applied research towards a quality assessment system*. Amsterdam: Royal Academy of Arts and Sciences; 2002.

Shah S, Ward JE. Outcomes from NHMRC public health research project grants awarded in 1993. *Aust N Z J Public Health* 2001; **25**:556–60.

Shani S, Siebzeiner M, I, Luxenburg O, Shemer J. Setting priorities for the adoption of health technologies on a national level: the Israeli experience. *Health Policy* 2000; **54**:169–85.

Spaapen J, Sylvain C. *Societal quality of research – towards a method for the assessment of the potential value of research for society*. SPSG Concept Paper 13. London: Science Policy Support Group; 1994.

Stryer D, Tunis S, Hubbard H, Clancy C. The outcomes of outcomes and effectiveness research: impacts and lessons from the first decade. *Health Serv Res* 2000; **35** (5 Pt 1):977–93.

Weisbrod BA. Costs and benefits of medical research: a case study of poliomyelitis. *J Political Econ* 1971; **79**:527–44.

Wisely J. *National R&D programme in the area of mother and child: programme report*. London: NHS London Regional Office; 2001.

Wisely J. *Assessment of the benefits from the National R&D Programme on Primary/Secondary Care Interface*. Unpublished report. London: NHS London Regional Office; 2001.

Wooding S, Hanney S, Buxton M, Grant J. *The returns from arthritis research. Volume 1: approach, analysis and recommendations*. Cambridge: RAND Europe; 2004.

Yesudian CAK. *Policy research in India: the case of regulating private health providers*. Geneva: The Alliance for Health Policy and Systems Research; 2001.



## Appendix 3

# Template for HTA review of literature on assessing the impact of research programmes

**Paper:**

**Description of method or application or both:**

**1. Description of method/framework:**

*Conceptual framework:*

*Dimensions of payback:*

*Research methods:*

*Other points:*

**2. Tabulation of empirical studies:**

*Study detail (country/setting/sponsor):*

*Study objectives:*

*Subject of assessment:*

*Study design/approach (framework/dimensions of payback/interviews, questionnaires):*

*Type of research evidence:*

*Findings:*

*Conclusions:*

*Internal validity:*

*Research recommendations:*

*Authors' comments on findings/conclusions/limitations:*

*Our/reviewer comments:*

*Conflicts of interest?*

*Methods transparent?*

*Description of aspects/dimensions of impact assessment?*

*Selection of sample?*

*Findings – are they clearly presented?*

*Conclusions – based on findings?*

*Internal validity?*

*Other:*



## Appendix 4

### Review of empirical literature/studies (by approach)

A commentary/narrative review is now presented on the main empirical studies that are particularly relevant for the assessment of the UK HTA Programme. *Table 1* presented outline characteristics on the 41 empirical studies identified and also the five that were only methodology. Not all 41 are covered in the following narrative review. Reference citations relate to the reference list at the end of this Appendix (p. 106).

#### The 'payback' approach

In the list of 41 studies, the 'payback' approach has been the most frequently used conceptual or methodological approach, informing eight of the empirical studies.<sup>1-9</sup>

**Buxton and Hanney**<sup>2,8</sup> reported eight short case studies using the payback framework. These case studies varied in detail and context, and were used to test and demonstrate the payback methods, and they are not discussed in any detail in this review. However, these early applications did highlight the difficulties associated with quantifying payback across all categories, with 'heroic assumptions' used in one of the case studies to estimate payback in the area of health sector benefits, and no attempt made to quantify payback in the category of broader economic benefits.

**Buxton and Schneider**<sup>3</sup> reported an application of the payback methods in Canada, in the assessment of payback from AHFMR-funded research. Although the study offers an empirical contribution, its main aim was methodological. This objective of the study was to test the payback model in the setting of AHFMR, not to assess the overall payback from AHFMR-funded research, nor to 'audit' individual projects" (p. 10). The study considered two projects funded by the AHFMR (external research), two health technology assessment reports/projects completed by the internal team at AHFMR and four programmes of research (lead researchers) funded by the AHFMR as capacity development. It used a 'desk' review of available information and materials on file, followed by interviews with key researchers. No questionnaires were used in this study. The

categories of payback that were used for the study are not explicitly stated, and the main focus of the study was on the sequence of research, and how the payback approach could be applied in the setting of AHFMR.

'On file' materials comprised research applications, reviewers' comments, annual reports, main publications and secondary accounts of research. Interviews for the health services research (HSR) projects and health technology assessment reports were undertaken to verify conclusions from on file materials, and to identify other categories of payback. For the research areas in basic/scientific research, a common structure of broad questions was used (structure presented in report), although interviews were individualised.

The findings stated by the authors are of a methodological nature, not directly/predominantly related to the measurement of impact. The authors state that the HSR case studies clearly show that the payback approach works in the context of AHFMR (Canada). The report presents a summary of the eight specific subjects of assessment (research areas/activities). Each of the case studies is presented against the payback research sequence headings (research needs assessment, interface, inputs, processes, primary and secondary outputs, applications and impacts or final outcomes). The impact assessment is covered in a very summary format under primary and secondary outputs and impacts or final outcomes. Each of the case studies is presented in a positive manner with evidence of impact or potential impact. However, the authors use a selective case study approach, and they state that the examples used were of a positive and convenient nature. The research areas/projects selected are not discussed in any detail, but the study states it selected case studies that were "good examples of applied HSR funded by the AHFMR ... that had already been fairly well documented". The two health technology assessment reports were selected based on information and researcher availability. For the examination of the basic/scientific research element of AHFMR (research capacity), the sample comprised four 'distinguished researchers', each being leaders in their field with research of

the highest standard. Although the authors state they are testing the method, it is never stated to what extent the method was being applied. The categories of the payback approach are not used to categorise payback from AHFMR programmes/projects, and where there is discussion of payback it is often speculative/expected impact rather than evidence based on the actual impact of research.

**Buxton and colleagues**<sup>9</sup> reported an application of the payback approach in an assessment of the payback from the research undertaken by the UK NHS R&D Programme of the North Thames region. The study sent questionnaires to lead applicants of all projects funded by the programme during the 1990s and was completed by 1997. A total of 164 questionnaires were sent out and 115 (70%) were returned. A sample of the projects was selected for further assessment as case studies.

The authors report that the questionnaires indicated payback from research against knowledge and benefits to future research and research use (the first two categories of payback), with an average of 2.2 publications (of any sort) per project, with just under 50% of these being journal articles (from 55 projects). The study included assessment of journal impact factor analysis and the authors state that they judged that the impact factor analysis may have underestimated the payback from some applied projects. Over one-quarter of projects reported contributing to postgraduate qualifications, but Buxton and colleagues had difficulty assessing the precise nature of the contribution of the North Thames projects to the qualifications reported. In terms of policy, practice and health benefits, the study reports that, considering each category separately, approximately 40% of respondents indicated their research had made an impact in one or more of these areas, and over 60% expected it do so. The study reports that over 80% of respondents indicated that they had an actual or expected benefit under one of the three categories.

Buxton and colleagues devised a scoring system to provide a benefit score for the projects. The paper does not provide detail on the scoring methods [it refers the reader to Hanney *et al.* (1999)<sup>10</sup> for detail] but it does report that the benefit scores derived by the impact project (via subjective assessment of reported payback) did not correlate well with impact findings derived from the use of bibliometric methods (i.e. publications, impact

factor analysis). Scoring indicated some impact even where the projects had not produced journal publications.

The case studies undertaken showed a general trend for questionnaires to understate the payback from projects, although there were cases in either direction. The authors suggest that although case studies may offer a more comprehensive and more reliable basis for impact assessment, a well-designed questionnaire might be a useful and sufficiently reliable technique to obtain a broad view of impact from research projects.

As with many of the impact studies identified in the literature search, this study may be subject to some bias due to the nature of the funding (via NHS Executive North Thames) and the close involvement of the project team with the funding body. However, it establishes some quantification of payback from the research funded by the UK North Thames R&D Programme, albeit self-reported payback in a number of cases (i.e. policy, practice and health benefits).

**Hanney and colleagues**<sup>4</sup> applied the payback approach to assess the impact of projects funded by the UK NHS R&D Implementation Methods Programme that was subsequently administered by the National Coordinating Centre for Service Delivery and Organisation (NCCSDO). The study had a broader remit than impact assessment alone (e.g. commissioning process, communication strategy). The study assessed the impact of funding for 36 research projects, using questionnaires to all lead researchers, questionnaires to users (using both electronic and postal format), documentary analysis and interviews with research commissioners. The response rate to lead researcher questionnaires was 83%, although some were only partly completed. The study collected views from two user groups, one via postal questionnaire (practitioners and researchers in the maternity field) with a 44% response rate. The other was via an electronic network (CHAIN) – but there were only 22 responses from 535 approached.

This study reports payback in terms of knowledge production and against various items related to the capacity to conduct further research. Around 120 publications were associated with the programme funding, 59 of which were in peer-reviewed journals. Citation analysis was undertaken (using ISI) – 40 journal articles in journals from the ISI database – but many articles were recently published and this reduced the value

of citation analysis. Lead researchers reported that 92 presentations had been made to academic audiences and 104 to practitioner/service groups (40 of these presentations were related to one project and 20 were on another individual project). Funding for a 15 further projects could potentially be associated with NCCSDO programme funded research, and at least nine projects reported benefits in terms of research training (four PhDs, three MDs). The study (e.g. questionnaire) considered the possible impact on health policy (as broadly defined) and practice, but did not attempt to consider final outcomes. Nine lead researchers reported that the project had already had some impact on policy, and 16 expected the project undertaken to have an impact on policy. For impact on practice, eight and 17 respondents indicated an impact already or an expected impact, respectively. The study also reports information available from within the programme on scores given by reviewers on the applications for funding and scores given by reviewers of the final outputs from the funded projects (this was undertaken for 19 of 36 projects).

The authors conclude that it was difficult to assess payback in a programme of research that was very broad. A degree of payback was reported, but it is not clear whether the payback analysis was altogether successful in this instance – with a broad range of projects, and a broadly specified project.

**Wooding and colleagues**<sup>7</sup> reported an impact assessment project using the payback approach. The project was undertaken for, and funded by, the Arthritis Research Campaign (ARC), and its main objective was to develop a system for evaluating arthritis research, for future use by the ARC. The study applied the payback framework, using interviews and documentary analysis, in the conduct of detailed case studies. No questionnaires were used in this application of the payback approach. The study conducted case studies on 16 research grants. The case studies were selected from 556 funded research grants awarded between 1990 and 1994 (a selection matrix was used – but essentially purposive). Information was gathered from documents and literature reviews, then semi-structured key informant interviews and bibliometric analysis was undertaken. The payback approach was used as a common structure for the case studies. The case studies also included a subjective scoring approach – with scoring systems developed by, and scores awarded by, the study authors.

The study concludes that there was strong evidence that there was a considerable range of research paybacks, and that the paybacks would not have been identified without the structured approach taken. The authors report payback from the funded grants against knowledge (302 papers, with 975 citations per year), research targeting and capacity building (28 PhDs/MDs, plus development of technological know-how, etc.), and in the informing of policy (informed recommendations and clinical guidelines). The study discusses difficulties identifying and assessing health sector benefits and economic benefits. Payback against these categories is inferred, with a presumption of benefits in terms of quality of life improvements for people with rheumatoid arthritis, and unquantified economic returns.

The authors state that there was good evidence that the payback method could be adapted for the ARC research programme. They found that the payback framework proved to be effective in capturing the diverse range of research outputs and outcomes. They also noted that the project grants tended to provide value for money because the payback appeared to be similar to that arising from other modes of funding that were better resourced.

**Wisely**<sup>5</sup> reported a study to assess the impact of the UK NHS R&D Programme on Mother and Child Health, which had commissioned and managed 51 projects (total £7.3 million) in 21 research areas (related to mother and child care). The study objectives were broader than an assessment of payback, and it reports details of project performance (funding and delivery of report). The study just used questionnaires to assess the impact of funded projects: no interviews or documentary analysis are reported (other than some *ex ante* and *post* project grading). The questionnaire used was adapted from that used by Buxton and colleagues<sup>9</sup> in an earlier payback study. Questionnaires were sent to 39 lead investigators of completed projects (39 of 51 projects completed), and there was a 67% ( $n = 26$ ) response rate. In questionnaire response data, 22 projects claimed to have published their findings, with two of the other four studies stating that they had submitted papers for publication. There were 54 journal articles published (the funding body was acknowledged in 35); results are not presented by study. All but one of the projects reported dissemination activity in the form of presentations, workshops and seminars. Eight projects reported benefits in terms of additional qualifications gained, with considerable input from the projects

in eight PhD awards (or equivalent). Nine projects reported payback from funding in terms of generation of additional future project funding (total about £2 million; with £1.3 million of this from one project), and 20 stated that the research had contributed to further research conducted by team members or others. Lead investigators for seven projects reported findings to have been used in policy/decision-making, and 13 projects expected findings to be used in policy/decision-making. Change in the behaviour of practitioners/managers was reported by eight projects, and eight expected their findings to influence practitioner behaviour in the future. Health service benefits had arisen from the projects in six instances (as reported by lead investigators), and 11 other projects expected benefits in the future.

The above results of the impact assessment are presented as self-report data from lead investigators, with some examples of health service benefits discussed in brief statements, but the authors also comment that many respondents described ways in which they (the lead investigators) hoped the findings may be used. The study authors indicate that research leads (teams) may not be the best placed to judge impact of projects in terms of the use of the research findings.

**Wisely**<sup>6</sup> reported a study to assess the impact of the UK NHS R&D Programme on Primary and Secondary Care Interface, which had commissioned and managed 70 projects (total £8.8 million). The study objectives were broader than an assessment of payback, and it reports details of project performance (funding and delivery of report). The study used questionnaires only to assess the impact of funded projects; no interviews or documentary analysis are reported (other than some *ex ante* and *post* project grading). The questionnaire used was adapted from that used by Buxton and colleagues<sup>9</sup> in an earlier payback study. Questionnaires were sent to 63 lead investigators (63 of 70 projects completed by then), and there was a 63% ( $n = 40$ ) response rate.

In questionnaire response data, 35 projects reported having published their findings, with the majority of the remaining five projects stating that they had submitted papers for publication. There were 89 journal articles published (the funding body was acknowledged in 57); results are not presented by study. All but one of the projects reported dissemination activity in the form of presentations, workshops and seminars. Nine

projects reported benefits in terms of additional qualifications gained, with considerable input from the projects in six PhD awards (or equivalent) and three MSc awards. Nineteen projects reported payback from funding in terms of generation of additional future project funding (total about £3.1 million), and 25 stated that the research had contributed to further research conducted by team members or others. Lead investigators for 14 projects reported findings to have been used in policy/decision-making, and a further 14 projects expected findings to be used in policy/decision-making. Change in the behaviour of practitioners/managers was reported by 11 projects, and 14 expected their findings to influence practitioner behaviour in the future. Health service benefits had arisen from the projects in 10 instances (as reported by lead investigators), and 15 other projects expected benefits in the future. The above results of the impact assessment are presented as self-report data from lead investigators; some examples of health service benefits are discussed in more detail by the authors.

### **Monetary value approach to estimating returns from research (cost–benefit analysis, or estimated cost savings)**

**Drummond and colleagues.**<sup>11</sup> The methodology was developed in this study to allow prospective assessment of the likely impact from proposed research (which is excluded from our review) but the application was to a completed programme of work in the form of a major clinical trial in the USA, the Diabetic Retinopathy Study, funded by the National Eye Institute from 1972 to 1981. The study adopted a decision theoretical approach and assessed the social costs and benefits of the research. Expert opinion was combined with survey data to derive estimates of the level of pre-existing practice and the likely diffusion rate of the trial results (which were known). A broad societal viewpoint was adopted for costs of the trial and treatment, untreated disease and medical assistance, rehabilitation services and lost production due to vision loss. It therefore considered both direct and indirect cost savings and uses the human capital approach to value health improvement.

Including the costs of lost production, the best estimate of net savings to society from the clinical trial was US\$2816 million over 22 years' use of photocoagulation therapy. This total reflected

\$2249 million savings to the government and \$1339 million to patients, but \$772 million costs incurred by third-party payers for increased level of photocoagulation therapy. Even excluding the costs of lost production there would be a net saving to society of \$231 million and a net gain of 279,000 vision years. The benefits came from cessation of less effective therapy and encouragement of effective therapy. The results were particularly sensitive to assumptions about the likely impact of the trial on clinical practice and on the probability of being treated. The conclusion was that the Diabetic Retinopathy Study was cost saving under many assumptions: "This study demonstrates that retrospective assessment of the potential net social benefit from clinical trials can be undertaken using the techniques of economic evaluation."

The authors were aware of some of the difficulties with such studies, including the fact that economists are divided on the issue of the relevance of including production losses and that the results were affected by the assumptions about likely impact of the trial on clinical practice.

This is a major study in the field of assessing the impact of medical research, but it is at the margins of inclusion in our review because the research study is not a full programme and the impacts past and future were estimated based on expert opinion.

**NIH.**<sup>12</sup> In two phases, 34 case studies of the impact of NIH-funded research were provided by their institutes, centres and divisions using a centrally provided framework. For each example, the reported NIH research input included one or more clinical trials and frequently several years of applied research. Basic research funds were excluded as were funds from other sources that might have supported the specific NIH research.

The approach assumed the NIH-funded work would impact on clinician behaviour and that estimates of the impact could be used to calculate a monetary figure for the cost savings. All benefits were combined into a single figure (or range) of monetary benefits: "The estimated savings are based on the difference between estimated direct plus indirect costs for a particular disease or condition before and after the innovation." Direct costs include costs of medical resources consumed in providing required healthcare and also non-medical costs such as custodial care. Indirect costs represent the productivity lost to society as a result of premature mortality or lost work days due to

morbidity. In this human capital approach, such costs are valued in terms of lost earnings.

Desk analysis was used to make the calculations of the cost savings which (apart from one example) are prospective savings for the cohort of patients who are expected to initiate treatment during 1 year. They would be replicated for future years. The annual savings reflected experts' judgments regarding the probable rate of adoption of the healthcare innovation. The estimates adjust for uncertainty in various parameters, including adoption rates.

In total, the research for the 34 examples cost, in 1992 prices, \$4339 million and produced annual cost savings estimated to be between a low of \$9318 and a high of \$13,554. The percentage benefit was greatest in diagnosis/screening and least for prevention. This led to the conclusion that: "The impact of these advances on improvements in quality of life and reductions in morbidity and premature mortality yields significant potential cost savings."

The process is explained in detail (in the available files, but not in this publication), but it is not always clear on what basis experts made their estimates for future implementation of findings. The calculations are detailed, but the results depend critically on the estimates of levels of implementation, expressed as ranges, and which might be seen as rather optimistic. Although there is mention of the role of research from elsewhere in the more detailed accounts, this does not appear in this publication and other subsequent research at HERG has shown how some of the NIH claims for the role of NIH research, even in the USA, are exaggerated.<sup>13</sup> Furthermore, there are also reservations about the figures produced by the human capital approach.<sup>14</sup> The study is based on selected examples from the whole NIH portfolio over many years, rather than a review of a specific programme. It does, however, provide useful examples of the large potential benefits.

**Raiten and Berman.**<sup>15</sup> This study was conducted in the USA by the Federation of American Societies for Experimental Biology (FASEB) and studied the benefits accruing in the USA, as a result of UK research into monoclonal antibodies (MAbs) that was originally conducted at Cambridge, UK. The aim was to address the increasingly important public policy issue of 'do the benefits of basic biomedical research justify the costs?', by developing a framework for conducting

analyses of the costs, benefits and impacts of basic biomedical research and undertaking a case study.

The study consisted of historical tracing of the research that led to the key findings, a CBA of the topic in general and a detailed study of one application of MAbs, namely HIV testing of the blood supply (this was incidentally one of the examples in the NIH study above). The benefits were measured in terms of (1) income losses avoided from lower production as a result of lost days of work and lost production or work due to accelerated mortality; (2) the reduction in medical care costs which would otherwise be incurred as a result of blood supply transmission of the HIV and subsequent development of AIDS; and (3) as secondary benefits, the output and employment of supporting manufacturing and services.

It is claimed that the historical tracing, beginning with the early developments in immunology and culminating in the hybridoma technology described by Kohler and Milstein in 1975,<sup>16</sup> provides corroboration of the length of the innovation process. The tracing also provides evidence of the inextricable and unpredictable role of non-directed, investigator-initiated fundamental research in the subsequent evolution of new technology. The economic analysis provides insights into the extent of the industry which has developed since 1975. A substantial overall return on investment is demonstrated. In the case study of a single application of MAb technology, the benefit to cost ratio for the initial investment in the development of a screening test for HIV containment of the blood supply is estimated to be 19:1.

The conclusion is drawn that: “This study establishes the utility of the combination of historical tracing, documents the important role of investigator-initiated research, and cost–benefit analysis to account for the return on the public’s investment.”

The focus of this important study is on basic research and thus a long way from health technology assessments. Furthermore, it looks at a stream of research rather than a specific programme. The specific case study on the work on MAbs is held up in UK science policy circles as a classic example of how key basic research was conducted in the UK, but the benefits mostly accrued elsewhere. It is not clear how the return on investment is calculated, especially given this history. It is an interesting development of CBA beyond the benefits usually given a

monetary valuation in assessments such as those by NIH.

**Weisbrod.**<sup>17</sup> This study examined the benefits from the Salk and Sabin research in the USA that led to the development of vaccines against polio. The aims of this early study were to examine the costs and benefits analysis of a medical success with the hope that “if other case studies of medical research are undertaken, we may someday be able to develop generalizations as to the nature of the probability distribution of rates of return from various types of medical research and application programs”. Second, Weisbrod hoped that this case study would help to emphasise the relationship between research and its application and the relevance of both to assessment of the social profitability of discovering new knowledge.

The benefits are worked out in monetary terms through valuing the health gain in terms of the sum of the market value of production lost because of premature mortality due to polio, the market value of production lost as a result of morbidity – illness and disability – caused by polio and the costs of resources devoted to treatment and rehabilitation of polio victims. The study involves estimating the time streams of research expenditure directed towards the disease (polio), the time streams of a number of forms of benefits resulting from (or predicted to result from) the application of the knowledge generated and the cost of applying that knowledge. Finally, internal rates of return on the research expenditures are computed using several alternative sets of assumptions regarding costs and benefits.

The findings were that, except under the most extreme assumptions, this research raised output and reduced treatment expenditures in amounts producing a rate of return on the research and application costs of at least 5%, or more probably 11–12%. A key conclusion was that analysis of the returns to medical research required the recognition of the inter-relatedness of the research itself and the procedures for applying the fruit of the research.

This is a classic and important example based on actual benefits from a programme of research, but has not led to the type of further studies that Weisbrod hoped for and recommended. The methods used included various alternative assumptions, and also the version of the journal article that appeared in his later book contains a postnote covering several points of clarification. However, criticisms of the human capital approach



are now greater and Weisbrod's paper is not even referenced in Murphy and Topel's 2003 major study of economic value of medical research.<sup>18</sup>

## HTA policy impacts and cost savings in Quebec

**Jacob and Battista.**<sup>19</sup> Using an early version of the methodological approach, the authors drew on part of an independent evaluation, by Price Waterhouse, of the Quebec Council on Health Care Technology Assessment (CETS). A case study analysis was conducted of each of the 10 CETS reports to determine its impact on the resulting actions taken by a range of key actors, called **critical incidents**. Each report was scored for its influence on a scale of 0 to ++++. Forty-five interviews were conducted with scientific and political partners and CETS members and documentary analysis undertaken. Then the Evaluation Branch of the Ministry examined the case studies to determine the financial impact of CETS, with savings defined as the net reduction in financial needs for the sectors covered by the reports. Savings do not necessarily represent a net reduction in total healthcare costs; typically, these financial gains resulted from limiting financial transfers between healthcare sectors.

The findings were that CETS performed well, with excellent quality of reports and pertinent topics. Eight of the 10 reports "influenced healthcare technology decisions, although not always to the same degree": two with ++++, five with ++ and one with +. Two case studies were described. Total savings were Can\$25 million, half of which represents yearly operating costs.

The Ministry felt that CETS reports contributed to improving the efficiency of decisions regarding health technology and represented a good investment for the healthcare system. Factors linked to this positive conclusion were identified. Conditions favourable to HTAs in Quebec include "a general receptivity to rationality in decision making ... an increasingly shared vision ... regarding the inevitability of choices to be made sets the stage for an effective use of information ... the healthcare system in Quebec is organised in such a way that information produced by the council can filter easily into the decision-making process that will lead to activation of mechanisms of regulation in the system in pursuing greater efficiency". Other important factors behind success were credibility, both scientific through rigorous standards of methodology in the health technology

assessments and politically because CET members were independent from the Ministry and not chosen on the basis of organisational affiliation, and local relevance. "The case studies were informative and revealed the dynamics of influence of each report." However, some were conducted just months after release, which was too soon for full effect, and there was no assurance that identified effects will last. In many cases, CETS reports were not the unique source of influence, therefore evaluating precise effects is "a difficult and delicate task. Appreciating the reports' impact boils down to deciding what would have happened if the reports had not been produced. The foundations of such an hypothesis rest on very soft ground".

This study of health technology assessments is included in the list from Gerhardus and Dintsios.<sup>20</sup> As with a later study by Jacob and McGregor,<sup>21</sup> this is an important approach and findings of substantial impact, but it is worth noting the acknowledged limitations, plus the special factors in the context that enhanced the role of receptors. There are also questions whether such cost saving figures are realistic in that it might be unlikely that the amount of cost savings would actually always have been available for additional expenditure by the healthcare sectors.

**Jacob and McGregor.**<sup>21</sup> This study was again conducted for CETS and developed the approach described above. In this case it focused on the 21 reports from CETS which had been in circulation long enough for at least some impact to be estimated (but it is not clear how far it includes some of the reports covered in Jacob and Battista's article from 1993<sup>19</sup>). Again, a wide range of techniques was used: analysis of documents and data banks of utilisation trends, interviews, surveys and desk analysis to estimate the cost savings.

The findings were that 18 of the 21 reports had an influence on policy, eight at the highest level. Some reports were based on the same topics and, therefore, counting topics, for 12 out of 16 evidence was found of considerable impact. The number of critical incidents for each level was also listed, with most being at level 1, but some at each level. The seven reports where the objective was cost minimisation (without reducing benefits or services) resulted in annual savings with a lower estimate of Can\$16 million and an upper estimate of \$27 million. Details of some case studies were given.

In its conclusions, this study demonstrates the feasibility of and limitations of assessing the effect

of health technology assessments on regulations: “In evaluating impact on policy, the method of identification of critical incidents, the categorisation of healthcare policies and the systematic use of documentation allows a certain degree of objectivity into the process. However, these methods have obvious limitations and depend heavily on the judgment of the analyst. The confirmation of impact by key actors is therefore important.” The conclusions noted some limitations of the methods. Detailed points include the fact that some groups who use the health technology assessment may prefer not to make an explicit reference to its impact on policy in documents, and other forms of influence such as on teaching materials were not looked for. In many cases health technology assessments were not the only source of influence. “Eventually, evaluating impact boils down to deciding what would have happened if the health technology assessment had not been produced.” The best insurance for impact is a request by a decider that an evaluation be made. There might be time lags before health technology assessment has an impact and this makes it harder to trace the impact.

This is a very important example of a methodological approach related to several dimensions of payback and with a detailed approach to scoring or scaling projects, and one showing high payback/impact from health technology assessments. There is a question as to how far it can actually be assumed that the money ‘saved’ would all have been available to be spent had the reports not been conducted and therefore whether these are all real savings. There is also a question about how far it could be generally applicable. One related issue might be that a sophisticated scaling or scoring system was developed and applied but it did not seem then to be used, for example, to compare average impact of different types of reports. The clear assumption that there would be impact leads them to analyse in discussion why three did not have any impact. The reasons for considerable impact of CETS is analysed in the example above.

**McGregor and Brophy.**<sup>22</sup> This report is included in this grouping because of the involvement of one of the authors of the above study. It is based on a hospital system within Quebec and it focuses on impact on policy and cost savings. The study is based on an evaluation of the work of a health technology assessment unit established in the McGill University Health Centre, Montreal, Quebec. The unit consists of a small technical staff

who access and synthesise the evidence incorporating local health and economic data, and a Policy Committee that makes recommendations. The techniques used in the evaluation are not described but it focused on the impact on the policy of the Health Centre and the budget. The impact of the first 16 studies since 2001 is reported.

The findings are that the recommendations from all 16 reports have been accepted. The recommendations fell into three categories: that the technology should be used, that it was not recommended or that its use should be more strictly limited than estimated demand. The potential benefit of introducing those technologies that were recommended are listed, but there seems to be no attempt to estimate actual health gain. An estimate was made, however, of the budget impact of following the recommendations. Some were estimated to increase costs, but 10 were estimated to produce savings.

The authors suggest that probable reasons for the success of this local in-house health technology assessment agency are “(i) relevance – selection of topics by administration with on-site production of health technology assessment allowing them to incorporate local data and reflect local needs; (ii) timeliness; and (iii) formulation of policy reflecting community values by local representative committee” (p. 263).

This study is another positive example from Quebec and again provides sensible reasons for the success of a health technology assessment agency. In this instance, however, no details are supplied as to how the impact analysis was conducted and the example possibly falls more into the category of a health technology appraisal agency than an assessment agency.

### **The levels of impact model developed by the Agency for Healthcare and Research Quality**

**Stryer and colleagues.**<sup>23</sup> The conceptual approach described was developed by the AHRQ as part of the process of analysing the data collected for its assessment of its Outcomes and Effectiveness Research (OER) programme. All 91 principal investigators in the OER funded between 1989 and 1997 were surveyed and asked about a range of impacts; 61 of the 91 reported and between them they covered 64% of the studies.

The findings were that OER has built a solid foundation for future quality improvement efforts by identifying problems, generating hypotheses, and developing new methodologies, but has had limited impact on healthcare policies, practices and outcomes.

The authors concluded that OER has had moderate but significant success meeting initial expectations for the field. Challenges for the next generation of OER include advancing from hypothesis generation to definitive studies of effectiveness, and acceleration of the process by which findings effect policy, practice and outcomes. The framework does not capture a number of important more general contributions of AHRQ's programme in OER: for example, it has played a major role in the expansion of interest and capacity in outcomes research in both the public and private sectors.

The particular importance of the study is that it produced a useful model and wide-ranging assessment of the impacts from a programme of research. The model came after data collection, and yet the paper reports it still did not capture all the benefits. The levels tend to reinforce our use of the stages of the payback model, even though evidence of impact above level 1 is only limited. There are question marks about representing the levels as a triangle because impact on policy is shown as being much broader than the impact on healthcare, and outcomes appear narrowly at the apex.

### **The Knowledge Utilisation model developed in Alberta**

**AHFMR.**<sup>24</sup> This study aimed to build on the study conducted the previous year of the HTA programme of the AHFMR. It was conducted by consultants for the AHFMR. The knowledge utilisation model was developed with the aim of applying a more extensive impact assessment to the 2001–2 products from the health technology assessment unit. It aimed to determine the degree to which the questions asked by requesters of health technology assessment products were answered, discover how much each product was used and by whom and learn if there were unexpected outcomes that came about as a direct result of the involvement of the health technology assessment unit with a requester's organisation. All the products that fitted the criteria (produced in 2001–2, publicly released and with an identifiable requester) were included. Using a structured

survey, nine interviews were conducted with requesters and/or users, out of 10 sought, and also three with authors of the HTA products interviewed.

The findings were that clients requested HTAs for various reasons, including information about efficacy and use, scope and generalisability of application and to support decisions. Most clients knew what they wanted but the health technology assessment unit was helpful when necessary. Overall clients were satisfied with the products and generally used the products to inform decision-making. Various supporting quotations were given in the report. In some cases some of the information provided an update for which there was no immediate use but it provided reassurance. Some clients found conceptual uses for the reports and used them to help establish research questions and advance evaluation agendas. Time was seen as the greatest barrier to use of the reports by health technology assessment staff. The report concluded that confidence in the information provided is the main reason for client satisfaction and most studies being used by clients.

The authors suggested that the questionnaire used for gaining information from requesters could be developed so that questions, such as 'do you use the health technology assessment product', could be accompanied by a list of potential responses. From this more directive, efficient instrument, a longitudinal database could be developed. Limitations included possible interviewee recall bias and positive response bias, requesters did not have much to offer, interviewee availability was limited, although the analysis attempted to cover 'utilisation' in a broad sense, it was difficult for clients to answer about this wider perspective, and the questions in the interview guides were not sensitive to the unit's definition of 'impact'. Although the study deliberately moved from the more restricted view of impact contained in the 2000–1 assessment towards a more multi-dimensional categorisation, it still proved impossible in this study to identify users beyond the immediate requesters.

### **The non-academic impact of socio-economic research**

**Molas-Gallert and colleagues.**<sup>25</sup> This paper describes the application of the model to one of two programmes, the AIDS programme, identified by the sponsor of the study, the ESRC in the UK, as being appropriate programmes on which to test

the conceptual approach being developed for them to assess the impact of their research on non-academic audiences. The ESRC's programme (1988–93) was on the social aspects of AIDS and consisted of 15 projects. (The other programme assessed was described in another paper but was not in the health field.) Interviews were the main research tool. One research team had not completed the report and one did two reports, therefore there were 13 sets of interviews with researchers and users and 43 in total.

The findings were that only a minority of researchers thought benefits from the projects included learning new techniques or the development of new ideas, but more referred to gaining further research funding. Links between researchers and users were weak before the start of the programme and remained weak, but the programme had an effect of strengthening mainly existing links and the number of links did increase slightly. "In general, networking rated low among researchers as a benefit arising from participation in the programme." There was some employment of academic experts involved in the AIDS programme by users, either in jobs or contracts.

Careers of the researchers developed in many ways, most retaining some continuity in the AIDS field, but others had to re-focus. Almost half have since played advisory roles and about one-third have been involved in contract research. They overwhelmingly pointed out that they had applied to this work skills and knowledge gained from the project. Half of the researchers thought that the programme had provided non-academics with tools to solve problems and had been used to develop policies. About 75% of users "said they have used project results to justify, confirm and help develop new courses of action". Such generic assertions might prove over-optimistic (and probably more user interviews were conducted for projects where impact was identified), but it was sometimes difficult for researchers to know of the indirect ways through which research had an impact. There were five main types of impact for which examples were collected: direct application with uncertain impact, for example a training book was produced, but use was restricted; dissemination of results with likely indirect impact, for example designing curricula and campaigns; indirect impact through methodological development, for example showed could do research on hard to reach groups; using research to target policy actions, for example a regional health authority used the results to inform the design of AIDS prevention and treatment

programmes; and a role in the launch of specific policy actions.

As a methodological study, some important conclusions were made for how such studies should be conducted, including that an impact assessment will require at least a two- or three-stage process with researchers being contacted first and users later. There is considerable diversity in communication and, therefore, "aggregate data on ... the percentage of projects that have established stable links with users, are of limited value". The interviews were of great value: "The preferable way to determine and assess the existence and nature of the impacts outside the academic sphere of ESRC-funded programmes is through detailed, project-by-project qualitative analysis." "Waiting between one and two years after the conclusion of each project to conduct an assessment would provide a balance between obtaining some long-term perspective while minimising the problems of identifying and tracing researchers and users." Tracking the original researchers was sometimes difficult and even more so users, whose memories of the project were often tenuous. Eventually, most user interviewees agreed to take part and the interviews were successful, despite the often initial hesitancy.

The broad techniques to develop were set out in the earlier scoping exercise and tender document for this study. The methods helped produce a range of findings and led to important methodological conclusions/recommendations. How far the full conceptual framework described above was helpful is less clear, but it did point up the complexity.

## The assessment of societal impact of research in The Netherlands

**Spaapen and Sylvain.**<sup>26</sup> Having described in great detail their conceptual model developed for the Dutch Sector Council (advisory councils for research in The Netherlands), as discussed above, Spaapen and Sylvain then tested it in relation to the work of two health research centres in The Netherlands. The aim was "to develop a methodological framework through which scientific and technological research can be assessed with respect to its potential value for society". Various techniques were adopted in the pilot study. First, the relevant characteristics of the communication structure in a particular field were defined by charting the various types of output, media and demand using surveys and/or

consulting relevant actors. Bibliometrics was also used. Second, various quantitative and qualitative techniques (bibliometrics and expert opinions, etc.) were used to establish the nature of the relevant issues and discussions between science, technology and society.

The importance of the empirical study for our review lies more in how the findings were presented than in the actual findings because these were not really programmes of research but rather centres with multiple funding sources and the actual dimensions of payback covered are rather modest. The six items included cover written output (visibility and range), non-written output (media and mobility) and media (relevant journal set and journal). Of central importance is the development of societal quality research profiles to present the findings about these items. The profiles are presented as a form of what has been referred to in the HERG/RAND study<sup>7</sup> as a 'spidergram' or payback profile and this influenced the report in The Netherlands described above.<sup>27</sup>

### **Selected other studies that focus on the impacts from a range of projects in a programme**

**Ferguson and colleagues.**<sup>28</sup> This study assessed the overall benefits to the NHS of regionally funded reactive research programmes of what were originally two NHS regions in the UK, Northern and Yorkshire, that were then merged to form the Northern and Yorkshire NHS region. The research team consisted of independent researchers and staff in the relevant regional R&D office. Three broad programme areas were covered: biomedical, HSR and primary and community care (PCC). The 158 projects started after April 1991 and finished before March 1996 were included.

The report states: "Some national research has been carried out in this area in the field of health services research<sup>29</sup> and other government departments have commissioned reports. The present study is set within the context of such wider research, which has explored the conceptual issues in much greater detail than is attempted here. Indeed it is important to emphasise that the present study is not a research study. What it has attempted to do is to apply the framework and lessons learned from elsewhere to identify some of the benefits which have arisen from NHS R&D

funding in one region" (para 1.4). Benefits were viewed as being 'multi-dimensional', and the final classification of outputs was peer-reviewed publications; non-peer-reviewed publications; presentations/seminars/workshops; changes in individual practice; changes in NHS service delivery; and career development of researchers, including ability to attract further research funding. Questionnaires were sent to all named applicants on projects; 51% of questionnaires were returned, covering 75% of all projects. The data provided in the questionnaires were quantitative and qualitative, but were all analysed using a coding frame, although the qualitative information was also useful in drawing attention to some important 'case studies', which were briefly written up in the report on the basis of information supplied by the survey. The desk analysis included the application of journal impact factors.

The findings were that 119 projects produced 230 peer-reviewed publications (average of 1.93), but 55 of the 119 produced no peer-reviewed publications. Most publications were in the top two quartiles of journal impact factors for their subject category. About 25% of projects claimed a change in NHS practice resulted from their project and led to more effective treatment, screening or patient management; 32% of projects made suggestions for changes in practice; 16% of HSR projects led to better informed commissioning or contracting; and 28% led to improved communication between professionals and with patients. The more detailed breakdown in the report shows that the 25% of projects claiming a change in NHS practice consisted of 35% of biomedical projects, 15% of HSR projects and only 5% of PCC projects.

Overall, 34% of projects led to increased research capacity and ability, and many thought that it had facilitated career progression. The total value of the grants included in the study was £2,232,405, but a further £6,277,322 had been obtained in research funding relating to the initial grant, representing a yield of 281%. Figures were given on various outputs for each of the three programmes and also for projects of different sizes. The authors' concluded: "in any sector of the economy a proportion of R&D expenditure will not pay off, but overall the results indicate a substantial return on regional investment in R&D projects. Across the whole range of outputs considered, the success rates appear impressive, in particular with regard to the ability of researchers to attract further research funding related to the

original award ... There is considerable evidence from our study to suggest that the return on investment from small projects can be substantial.”

The authors recognise that the study relies entirely on self-reporting, but note that more details of the qualitative data analysis are given in the full report. In some cases more than one researcher returned a questionnaire for a project and all questionnaires were analysed. “Although information was collected on conference presentations and non-peer reviewed output, this did not in general add to the descriptive statistical analysis.” Difficulties were encountered in collating a reliable, validated database of regionally funded R&D projects. There is no mention in the paper of any attempt to correlate factors, either at the overall level or in case studies that might be linked to impact, beyond noting differences between the programmes.

This study was partly informed by the HERG payback framework and was fairly broad but does not go as far as outcomes. The positive findings about the impact from the research are particularly important in two regards. First, this was reactive research. Second, the returns from small projects were sometimes substantial, which can be compared with some of the findings from Wooding and colleagues.<sup>7</sup>

**Hailey and colleagues.**<sup>30</sup> This study analysed the impact of the health technology assessment reports from the National Health Technology Advisory Panel (NHTAP) covering the 20 technologies assessed by 1988. A framework set out the various types of impact, with that on policy the main factor, but also how far the report was used as a source and as educational material, and some consideration of impact on practice. Very limited detail is given on methods beyond referring to comparing recommendations and assessments and policy activities and other events (possibly synthesising published data and expert opinion – unclear), and requests received for copies of reports and citations.

The findings were that for 11 of 20 technologies, “the NHTAP reports appear to have had a significant influence in the short to medium term, on the basis of major recommendations and subsequent government or other action”. In three there is probably short-term influence. Sixteen proved useful as a source and educational material. Most of the recommendations dealing with introduction/reimbursement for technologies were accepted, the outcomes on accreditation and

safety remained uncertain, and recommendations on research or further study led to significant evaluations in several areas. The authors concluded that the NHTAP’s assessments and recommendations “can have an important influence on policy and the subsequent use of a technology but that much depends on the presence of stable and receptive policy areas ... the level of impact is not necessarily related to the depth of the analysis – timing is crucial impact on immediate government funding and licensing decisions may be relatively quick. The impact of recommendations which affect professional bodies and concern matters such as preparation of guidelines may take longer to become apparent”.

This study is important as a pioneering study and the case studies provide some useful examples of impact, but are rather variable. An important point is that impact on policy can be quick. The study is at the margin of what should be included in our review because it is based more on NHTAP recommendations than specifically on a programme of research. The article provides evidence for the broader study of Australian health technology assessments from a range of bodies<sup>31</sup> and a later account of one of the case studies is given as a separate paper by Hailey and Crowe.<sup>32</sup>

**Hailey and colleagues.**<sup>33</sup> This study considered the impact of the 20 short, rapid, health technology assessment reports, ‘Technotes’, developed in a province where health ministry and authorities needed them quickly. The main focus was on whether the report had been helpful to the decision-making process. The people requesting the assessments, or who might be influenced by the findings, were approached. Face-to-face discussions and written requests for feedback were used, covering clarity and relevance of advice in assessment and whether helpful. Checks on the quality of the reports were made using, for different reports, subsequent refereed health technology assessment reports, monitoring literature and comments from experts in the field. No details were given on how many interviews were carried out.

The findings were that 14 assessments had an influence on policy and other decisions, as judged by responses from those who had requested advice. Another four provided guidance, while having less immediate influence, and two had no apparent impact. The quality of the assessments was considered acceptable, on the basis of the literature that subsequently became available and

from comments received. The authors concluded that Technotes were considered a useful component of a health technology assessment programme and contributed to policy and administrative decisions in many cases, often where there was an urgent need, but should be regarded as provisional appraisals and followed up with more detailed evaluation where possible.

The study is again part of the body of evidence that shows a high proportion of impact for some health technology assessments, in this case for direct and urgently requested reviews, but it is not clear how far this should be regarded as a programme of research.

**NHS Executive South West.**<sup>34</sup> The 162 completed projects funded by the region between 1991 and 1997 were examined using an end of project assessment form including questions on publications, dissemination, higher degrees, other outputs and further research. The form was distributed by the regional office; 109 questionnaires were returned by lead researchers (74%).

The findings were that 55 projects reported at least one peer-reviewed paper (with non-RCTs least likely to publish in peer-reviewed journals, and laboratory studies and quantitative surveys most productive). Thirteen of the 55 had no impact factor, and only two papers had an impact factor >5; the mean number of citations was 1.95 but most were too recent for any citations at all. A total of 39 projects led to higher degrees and 73 projects reported active dissemination. The authors noted the time lags in the production of outputs and a limitation on analysis of outputs is the reliance on researchers to provide up-to-date information on publications. They noted there are debates about the limitations of ISI's impact factors, especially low coverage of some types of research.

Most of the report deals with processes, not impacts, and even those impacts considered were really just outputs but the assessment described here was leading to further work.

**Shah and Ward.**<sup>35</sup> The authors studied the outcomes from the research awards made in 1993 by the Public Health Research and Development Committee (PHRDC) of the National Health and Medical Research Council (NHMRC). They wanted to develop "a novel audit" (p. 556) of research and take it further than indicators of publications. They used a self-administered survey that asked about publications, research degrees

linked to the projects, whether the project had influenced first policy, and second practice to improve public health or delivery of health services. They also asked about study design. Surveys were sent to all chief investigators awarded new project grants (32) or continuing project grants (31) in 1993 following a telephone call to all contactable investigators before the questionnaire mail-out. There was a response rate of 69%, but some projects were removed from the total if their chief investigator could not be traced or had left Australia.

Desk analysis included using databases to check impact factors of journals and whether publications were on MEDLINE and/or EMBASE, and checking whether influence on policy, practice or both was associated with peer-reviewed publication in an Australian journal. Cell sizes prevented testing for associations between influence on policy, practice or both and journal impact factors.

The findings were that 82% of projects produced peer-reviewed articles, 68% gave presentations without published proceedings, 47% with published proceedings and various other publications and media outlets; 61% provided postgraduate research training: 25 PhDs, 14 MPH, eight MSc and one MD. About 58% strongly agreed or agreed their research had influenced policy to improve public health, 69% that it had influenced practice and 53% both. Influence on policy or practice or both was not associated with peer-reviewed articles in an Australian journal. The most highly rated strategy to enhance dissemination was greater demand for research results from among policy makers. The authors concluded that: "A pleasing proportion of projects funded by PHRDC in 1993 generated peer-reviewed publications and provided research training. Recipients perceived their research has influenced policy and practice. Policy makers emerge as a key target for training in research transfer." They also claim the findings "refute concerns that research is often represented as remote and elitist, driven by the interest of the investigator rather than the urgency and amenability of the problem to solution by research". Interest in research audits and development of valid outcome measures are increasing.

While the authors acknowledge methodological limitations, they refer only to the modest response rate and 1 year's grants, which means the generalisability is unclear. Nevertheless, they state,

“We offer our survey instrument as a prototype tool” (p. 560).

This is a most useful study that shows how surveys and bibliometric analysis can be used in a way similar to earlier ones by HERG to gather information on a range of paybacks from a specific research programme. It is also interesting that the impact on practice is claimed to be higher than on policy: this is contrary to Gerhardus’s conclusion from a review of health technology assessments<sup>20</sup> and it would have been useful to see the exact wording of questions on policy and practice. It has positive results, but is limited in that there was no attempt to check on self-reported data or explore them further.

**Shani and colleagues.**<sup>36</sup> This report studied the work of the Israeli Ministry of Health’s Medical Technologies Administration (MTA) in assessing a new list of technologies in 1999 for possible inclusion in the National List of Health Services (NLHS). The MTA, in cooperation with the Israeli Center for Technology Assessment in Health Care, conducted clinical, epidemiological and economic evaluations on 84 proposed technologies. The findings were fed into the Medical Technology Forum, which made recommendations on priorities to the National Advisory Committee. The study provides a commentary/account of methods of MTA and the influence of its work on the Forum and its impact on decision-making of the National Advisory Committee in relation to resource allocation in terms of which additional technologies should be made available. No account is given of the methods used in producing this commentary.

The main findings were that all the assessments fed into the Forum’s work and about 70% of the Forum’s recommendations were accepted immediately. The Committee then focused on the remaining 30% and made some changes in priority. Eventually, 50 were recommended and approved, 44 (88%) of which had been recommended for adoption by the Forum. The authors concluded that, “Unlike the 1997 update of the NLHS, which was described as being explicitly implicit, the 1999 update is an explicit process trying to blend healthcare professionals with lay participation. The process was widely accepted in Israel by government officials, healthcare professionals, politicians, and the courts.”

This study gives an interesting account of the effect of health technology assessments. It

provides a high level of impact and was included in the Gerhardus review. The study also makes international health technology assessment comparisons. It is possibly reasonable to consider the work of MTA as a programme, but this study is more of a commentary because it gives no details about its own methods and, furthermore, the evaluations are more like NICE TARs than the other research of the NHS HTA Programme. It is an example of where the system/institutional arrangements are well placed to allow health technology assessments to make an impact.

## References

1. Alberta Heritage Foundation for Medical Research (AHFMR). *Assessment of health research fund outputs and outcomes: 1995–2003*. Edmonton: AHFMR; 2003.
2. Buxton M, Hanney S. How can payback from health services research be assessed? *J Health Serv Res Policy* 1996;1:35–43.
3. Buxton M, Schneider W. Assessing the ‘payback’ from AHFMR-funded research. Edmonton: AHFMR; 1999.
4. Hanney S, Soper B, Buxton M. *Evaluation of the NHS R&D Methods Programme*. HERG Report 29. Uxbridge: Brunel University; 2003.
5. Wisely J. National R&D Programme in the area of mother and child: programme report. London: NHS London Regional Office; 2001.
6. Wisely J. Assessment of the benefits from the National R&D Programme on Primary/Secondary Care Interface. Unpublished Report. London: NHS London Regional Office; 2001.
7. Wooding S, Hanney S, Buxton M, Grant J. Payback arising from research funding: evaluation of the Arthritis Research Campaign. *Rheumatology* 2005; 44:1145–56.
8. Buxton M, Hanney S. *Assessing payback from Department of Health research and development: preliminary report. Volume 1: the main report*. Uxbridge: Health Economics Research Group, Brunel University; 1994.
9. Buxton M, Hanney S, Packwood T, Roberts S, Youll P. Assessing benefits from Department of Health and National Health Service research and development. *Public Money Manage* 2000;20:29–34.
10. Hanney S, Davies A, Buxton M. Assessing benefits from health research projects: can we use questionnaires instead of case studies? *Res Eval* 1999;8:189–99.
11. Drummond MF, Davies LM, Ferris FL III. Assessing the costs and benefits of medical research: the diabetic retinopathy study. *Soc Sci Med* 1992; 34:973–81.



12. National Institutes of Health. *Cost savings resulting from NIH research support*. Bethesda, MD: US Department of Health and Human Services, National Institutes of Health; 1993.
13. Hanney S, Mugford M, Grant J, Buxton M. Assessing the benefits of health research: lessons from research into the use of antenatal corticosteroids for the prevention of neonatal respiratory distress syndrome. *Soc Sci Med* 2005;**60**:937–47.
14. Buxton M, Hanney S, Jones T. Estimating the economic value to societies of the impact of health research: a critical review. *Bull World Health Organ* 2004;**82**:733–9.
15. Raiten D, Berman S. *Can the impact of basic biomedical research be measured? A case study approach*. Bethesda, MD: Life Sciences Research Office, Federation of American Societies for Experimental Biology; 1993.
16. Köhler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 2005;**256**:495–7.
17. Weisbrod BA. Costs and benefits of medical research: a case study of poliomyelitis. *J Political Econ* 1971;**79**:527–44.
18. Murphy K, Topel R. *Measuring the gains from medical research. An economic approach*. Chicago, IL: University of Chicago Press; 2003.
19. Jacob R, Battista R. Assessing technology assessment. *Int J Technol Assess Health Care* 1993;**9**:564–72.
20. Gerhardus A, Dintsios Ch. *The impact of HTA reports on health policy – a systematic review*. Cologne: German Agency of Health Technology Assessment at German Institute for Medical Documentation and Information (DAHTA) (DIMDI); 2004 (in German).
21. Jacob R, McGregor M. Assessing the impact of health technology assessment. *Int J Technol Assess Health Care* 1997;**13**:68–80.
22. McGregor M, Brophy M. End-user involvement in health technology assessment (HTA) development: a way to increase impact. *Int J Technol Assess Health Care* 2005;**21**:263–7.
23. Stryer D, Tunis S, Hubbard H, Clancy C. The outcomes of outcomes and effectiveness research: impacts and lessons from the first decade. *Health Serv Res* 2000;**35** (5 Pt 1):977–93.
24. Alberta Heritage Foundation for Medical Research (AHFMR). *A study of the impact of 2001–2002 health technology assessment products: final report*. Edmonton: AHFMR; 2003.
25. Molas-Gallart J, Tang P, Morrow S. Assessing the non-academic impact of grant-funded socio-economic research: results from a pilot study. *Res Eval* 2000;**9**:171–82.
26. Spaapen J, Sylvain C. Societal quality of research – towards a method for the assessment of the potential value of research for society. SPSC Concept Paper No. 13. London: Science Policy Support Group; 1994.
27. Royal Netherlands Academy of Arts and Sciences. *The societal impact of applied research towards a quality assessment system*. Amsterdam: Royal Academy of Arts and Sciences; 2002.
28. Ferguson B, Kelly P, Georgiou A, Barnes G, Sutherland B, Woodbridge B. Assessing payback from NHS reactive research programmes. *J Manage Med* 2000;**14**:25–36.
29. Hanney S, Buxton M, Brown J, Bryan S, Ratcliffe J, Walden L, et al. *Assessing payback from Department of Health research and development: second report: report to the Department of Health RandD Division: Vol. 2: ten further case studies*. HERG Research Report 24. Uxbridge; Brunel University; 1997.
30. Hailey DM, Cowley DE, Dankiw W. The impact of health technology assessment. *Commun Health Stud* 1990;**14**:223–34.
31. Hailey D. The influence of technology assessments by advisory bodies on health policy and practice. *Health Policy* 1993;**25**:243–53.
32. Hailey DM, Crowe BL. The influence of health technology assessment on the diffusion of MRI in Australia. *Int J Technol Assess Health Care* 1993;**9**:522–9.
33. Hailey D, Corabian P, Harstall C, Schneider W. The use and impact of rapid health technology assessments. *Int J Technol Assess Health Care* 2000;**16**:651–6.
34. NHS Executive South West. *An evaluation of the regional R&D project grants scheme*. Bristol: NHS Executive South West; 1999.
35. Shah S, Ward JE. Outcomes from NHMRC public health research project grants awarded in 1993. *Aust N Z J Public Health* 2001;**25**:556–60.
36. Shani S, Siebzehner M, Luxenburg IO, Shemer J. Setting priorities for the adoption of health technologies on a national level: the Israeli experience. *Health Policy* 2000;**54**:169–85.



## Appendix 5

### Total hits for HTA projects in survey

Type	Project code	Web downloads (records from April 2006)
PR	93/03/06	–
PR	93/03/26	11,674
PR	93/03/31	–
PR	93/06/04	–
PR	93/07/66	20,437
PR	93/07/68	13,735
PR	93/14/09	8,014
PR	93/17/11	10,291
PR	93/17/13	8,302
PR	93/17/43	12,517
PR	93/18/12	10,492
PR	93/18/26	–
PR	93/22/09	16,692
PR	93/23/19	8,488
PR	93/25/05	150,841
PR	93/26/16	27,608
PR	94/04/09	4,740
PR	94/14/20	6,527
PR	94/16/03	22,991
PR	94/18/05	22,466
SR	94/20/01	13,607
PR	94/22/24	11,035
PR	94/22/26	30,266
PR	94/24/03	5,746
PR	94/33/18	–
PR	94/39/14	19,756
PR	94/40/05	37,682
PR	94/40/27	27,597
PR	94/40/38	15,356
PR	94/41/17	–
PR	94/42/06	–
PR	94/43/04	39,623
PR	94/45/02	5,933
PR	94/46/01	–
PR	94/48/03	11,988
PR	95/06/18	24,500
PR	95/07/19	13,517
PR	95/08/05	6,796
PR	95/16/06	17,131
PR	95/17/06	18,301
PR	95/25/03	–
PR	95/30/02	18,589
PR	95/31/05	8,175
PR	95/33/01	–
PR	96/02/05	6,050
PR	96/04/06	13,323
PR	96/06/07	6,004
PR	96/07/03	–
PR	96/08/01	27,120
PR	96/09/14	–

*continued*

Type	Project code	Web downloads (records from April 2006)
PR	96/11/02	21,315
PR	96/12/26	21,321
PR	96/13/04	–
PR	96/15/05	35,014
PR	96/19/06	–
PR	96/22/11	4,635
PR	96/31/05	8,007
PR	96/32/08	8,430
PR	96/40/07	11,774
PR	96/40/15	7,994
PR	97/20/06	15,099
PR	98/16/04	4,785
SR	93/05/01	27,079
SR	93/10/01	10,184
SR	93/26/17	–
SR	93/29/01	52,374
SR	93/30/03	35,077
SR	93/33/03	34,173
SR	93/34/04	12,564
SR	93/35/02	9,220
SR	93/56/99	25,819
PR	94/05/03	11,138
SR	94/36/04	41,049
SR	95/09/03	31,833
SR	95/11/04	59,015
SR	95/14/01	21,135
SR	95/22/04	24,780
SR	95/29/04	17,407
SR	95/35/02	9,220
SR	95/36/02	19,104
SR	95/38/01	11,341
SR	95/40/01	14,328
SR	95/41/01	9,547
SR	95/55/03	22,313
SR	96/01/03	23,212
SR	96/10/03	49,009
SR	96/16/02	27,205
SR	96/17/01	100,257
SR	96/23/56	21,494
SR	96/33/02	9,890
SR	96/35/01	15,892
SR	96/37/01	18,635
SR	96/41/03	20,841
SR	96/45/02	23,353
SR	96/47/01	48,296
SR	97/07/01	63,767
SR	97/12/01	20,000
SR	97/13/04	21,219
SR	97/15/03	19,928
SR	97/23/02	45,192

*continued*

Type	Project code	Web downloads (records from April 2006)
SR	97/27/02	49,069
SR	97/30/99	24,839
SR	97/38/03	–
SR	98/02/02	24,528
SR	93/17/22	–
SR	98/29/04	9,375
SR	98/30/01	39,615
SR	98/36/02	–
SR	98/38/01	10,616
SR	99/21/02	22,929
SR	99/23/04	–
SR	99/29/02	2,977
SR	99/45/01	8,658
SR	01/05/02	–
SR	01/16/02	5,994
SR	99/02/02	101,211
HTA TAR	99/05/01	26,154
HTA TAR	99/05/02	33,485
HTA TAR	99/05/03	49,679
HTA TAR	99/05/04	18,052
HTA TAR	99/05/05	21,142
HTA TAR	99/09/50	27,785
HTA TAR	00/25/01	25,562
HTA TAR	01/28/01	24,278
HTA TAR	01/29/01	10,266
HTA TAR	01/30/01	17,872
HTA TAR	01/31/01	42,844
HTA TAR	01/32/01	27,031
HTA TAR	01/33/01	13,243
HTA TAR	01/34/01	55,967
HTA TAR	01/35/01	9,639
HTA TAR	01/36/01	29,338
HTA TAR	01/39/01	50,985
HTA TAR	01/63/01	20,704
HTA TAR	02/14/01	15,845
HTA TAR	02/15/01	11,876
HTA TAR	02/30/01	14,276
HTA TAR	02/43/01	22,377
NICE TAR	99/13/01	10,625
NICE TAR	99/15/01	18,747
NICE TAR	99/16/01	18,854
NICE TAR	99/17/01	19,538
NICE TAR	99/18/01	15,485
SR	95/09/03	31,833
NICE TAR	00/01/01	12,008
NICE TAR	00/02/01	15,599
NICE TAR	00/03/01	23,761
NICE TAR	00/04/01	11,502
NICE TAR	00/04/02	17,803
NICE TAR	00/05/01	18,387
NICE TAR	00/05/02	29,312
NICE TAR	00/06/01	17,214
NICE TAR	00/07/01	17,228
NICE TAR	00/08/01	34,196

Type	Project code	Web downloads (records from April 2006)
NICE TAR	00/09/01	8,740
NICE TAR	00/10/01	16,114
NICE TAR	00/11/01	10,165
NICE TAR	00/12/01	8,448
NICE TAR	00/12/02	8,666
NICE TAR	00/13/01	20,260
NICE TAR	00/14/01	39,087
NICE TAR	00/15/02	13,296
NICE TAR	00/15/03	8,535
NICE TAR	00/16/01	10,373
NICE TAR	00/17/01	25,015
NICE TAR	00/18/01	37,000
NICE TAR	00/18/02	11,459
NICE TAR	00/19/01	24,669
NICE TAR	00/20/01	52,102
NICE TAR	00/21/01	23,394
NICE TAR	00/21/02	14,027
NICE TAR	00/22/01	28,833
NICE TAR	00/23/01	20,588
NICE TAR	00/24/01	28,566
NICE TAR	01/20/01	16,531
NICE TAR	01/21/01	38,938
NICE TAR	01/22/01	23,489
NICE TAR	01/23/01	33,237
NICE TAR	01/24/01	29,916
NICE TAR	01/25/01	16,523
NICE TAR	01/26/01	21,896
NICE TAR	01/27/01	32,202
NICE TAR	01/46/01	20,595
NICE TAR	01/48/01	14,290
NICE TAR	01/49/01	15,499
NICE TAR	01/50/01	28,496
NICE TAR	01/50/02	295
NICE TAR	01/52/01	28,981
NICE TAR	01/53/01	19,927
NICE TAR	01/54/01	31,358
NICE TAR	01/55/01	32,265
NICE TAR	01/56/01	14,219
NICE TAR	01/57/01	25,083
NICE TAR	01/58/01	27,149
NICE TAR	01/59/01	9,348
NICE TAR	01/60/01	28,566
NICE TAR	01/61/01	28,726
NICE TAR	01/62/01	15,914
NICE TAR	02/16/01	41,728
NICE TAR	02/17/01	24,402
NICE TAR	02/18/01	14,914
NICE TAR	02/19/01	16,931
NICE TAR	02/20/01	21,115
NICE TAR	02/27/01	22,783

Mean number hits for all 204 projects = 20,440  
Total number hits = 4,169,802

PR, primary research; SR, secondary research.

## Appendix 6

### HTA reports 1997–2006 considered by the National Screening Committee (NSC)

Vol. (issue)	HTA publications	NSC action
10(14)	Cost-effectiveness of screening for oral cancer in primary care. April 2006	This report reinforced NSC policy not to screen
10(3)	Computed tomography screening for lung cancer. January 2006	This report reinforced NSC policy not to screen and led the NSC/MRC to commission review of evidence relating to breath test screening
10(11)	Screening for thrombophilia. May 2006	Policy position reaffirmed by the NSC
9(13)	Cervical screening programmes: can automation help? March 2005	The cancer screening programmes used this report
9(40)	Randomised control trial of targeted screening versus systematic population screening for atrial fibrillation in over 65s. October 2005	On the basis of this report the NSC maintained its position not to offer whole population screening for atrial fibrillation. This does not reduce the need to maintain clinical awareness
9(44)	Screening in first year of life for congenital heart disease. December 2005	This has been built into the protocol for the newborn physical examination. The HTA report supported the NSC policy of screening by auscultation and allowed the proposal to introduce screening by pulse oximetry to be turned down
8(12)	Clinical and cost-effectiveness of neonatal screening for inborn errors of metabolism by TMS. March 2004	The evidence appears to support the introduction of tandem mass spectrometry into a UK neonatal screening programme for PKU and MCAD deficiency combined. An evaluative study of MCAD screening has been set up based on the 2 HTA reports of 1997
8(33)	Psychosocial aspects of genetic screening of pregnant women and newborns : a systematic review. August 2004	This has informed our educational resources development for both women and professionals
7(10)	Evaluation of molecular tests for prenatal diagnosis of chromosome abnormalities. May 2003	This report has been the basis of the development of the policy for the diagnostic testing for Down's syndrome, a topic not covered by the SURUSS Report (see below)
7(31)	Lowering blood pressure to prevent myocardial infarction and stroke. November 2003	The conclusions of this report will be adopted by the Diabetes, Heart Disease and Stroke Prevention project set up by the NSC in October 2003 at nine sites in England
7(30)	Value of digital imaging in diabetic retinopathy. November 2003	The report concluded that digital imaging was effective and this is being carried out in the sight-threatening diabetic retinopathy screening programme throughout the UK
7(16)	Screening for Fragile X syndrome: a literature review and modelling study. August 2003	On the basis of the two reports the NSC decided that screening for Fragile X should not be offered in pregnancy, (see above)
7(11)	First and second trimester screening for Down's syndrome: the results of the Serum, Urine and Ultrasound Screening Study (SURUSS). May 2003	On the basis of this report and wide ranging discussion with experts, the NSC policy is to screen all women for Down's syndrome

*continued*

Vol. (issue)	HTA publications	NSC action
7(6)	The cost-effectiveness of screening for <i>Helicobacter pylori</i> to reduce mortality. (March 2003)	On the basis of this report the NSC has amended its policy guidance to emphasise that there should be neither screening for gastric cancer nor screening for <i>Helicobacter pylori</i>
6(31)	Screening for hepatitis C among injecting drug users and in GUM clinics. December 2002	The HTA report is clear that universal screening for hepatitis C should not be introduced in genitourinary medicine clinics. Therefore, the NSC policy will not be changed but will be reviewed in 2007
6(11)	Screening for gestational diabetes: a systematic review and economic evaluation. November 2002	On the basis of this report the NSC decided that screening for diabetes should not be offered in pregnancy
5(7)	As assessment of screening strategies for Fragile X syndrome in UK. March 2001	On the basis of the two reports the NSC decided that screening for Fragile X should not be offered in pregnancy (see below)
4(29)	Screening for hypocholesterolaemia. November 2000	On the basis of this report the NSC recommended a systematic approach to case finding in preference to population screening
4(18)	Liquid based cytology in cervical screening. August 2000	On the basis of this report and following pilots the NHS Cervical Screening programme will be converting to liquid-based cytology (LBC)
3(14)	A systematic review of the role of human papillomavirus testing within a cervical screening programme. 1999	On the basis of this report, the NSC recommended a pilot study to assess the practicability of using HPV testing to improve the management of women with equivocal smear results
3(11)	Antenatal and neonatal haemoglobinopathy screening in the UK: review and economic analysis. 1999 (two reports)	A national programme of screening for sickle cell disease and thalassaemia is being introduced on the basis of these reports
3(8)	Screening for cystic fibrosis. 1999	On the basis of this report, the feasibility of national roll-out of an antenatal screening programme, currently on offer only in Edinburgh and the Lothians, is being explored. Those commissioning health services were recommended not to fund neonatal screening for cystic fibrosis but the publication of a report from the major RCT of screening for cystic fibrosis in 2001 led to a decision to introduce neonatal screening
2(11)	Detection, adherence and control of hypertension for the prevention of stroke: a systematic review. 1998.	A second report on hypertension is expected and the NSC will consider these two reports in 2002
2(9)	Screening for speech and language delay: a systematic review of the literature. 1998	On the basis of this report, the NSC recommended that there should not be a national programme of screening for speech and language delay, principally because of the uncertainty about natural history
2(2)	Screening for ovarian cancer: a systematic review. 1998	On the basis of this review, the NSC supported an RCT of screening for ovarian cancer which was subsequently funded by the MRC
2(1)	Antenatal screening for Down's syndrome. 1998 Ultrasound screening in pregnancy: a systematic review of the clinical effectiveness, cost-effectiveness and women's views. 2000	The NSC recommended the introduction of a national programme to ensure that all pregnant women were offered a test and follow-up which met explicit quality criteria. This was linked to recommendations about ultrasound screening in pregnancy to minimise variations and quality problems

continued

Vol. (issue)	HTA publications	NSC action
I(10)	A critical review of the role of neonatal hearing screening in the detection of congenital hearing impairment. 1997	The NSC recommended the introduction of universal neonatal hearing screening using auto-acoustic technology to replace the distraction test currently offered by health visitors. Funds were allocated and screening is being introduced in a systematic way
I(8)	Pre-school vision screening. 1997	On the basis of this report, the NSC conducted a major review of vision screening and although there was no clear evidence of benefit decided to support the retention of a simplified service with better training of staff
I(7)	Neonatal screening for inborn errors of metabolism: cost, yield and outcome. 1997 (two reports)	On the basis of these two reports, the NSC's recommendation was that tandem mass spectrometry screening should not be introduced as a generic service but that further work should be done to review the evidence about the costs and benefits of screening for specific diseases. It has also set up a National Bloodspot Screening Programme Centre
I(4)	Screening for Fragile X syndrome. 1997	The NSC recommended that this should not be introduced
I(2)	Diagnosis, management and screening of early localised prostate cancer. 1997 (two reports)	On the basis of these reports, the Department of Health issued a letter stating that prostate cancer screening should not be offered until new evidence was available. Continuing debate based on these reports led to both the introduction of the Prostate Cancer Risk Management Programme, designed to improve PSA testing, and the commissioning of a major randomised controlled trial on different methods of treating early prostate cancer
Other	Study to evaluate the most cost effective way to screen for <i>Chlamydia</i>	Submitted to NSC but not published
In progress	<p>Screening for type 2 diabetes. Literature review</p> <p>Clinical and cost-effectiveness of screening for open-angle glaucoma</p> <p>Screening for early age-related macular degeneration</p> <p>Randomised trial of HPV virus testing in primary cervical screening</p> <p>Prenatal screening to prevent group B strep</p> <p>Antenatal screening for HbOs in primary care</p> <p>Accuracy and cost effectiveness of rapid diagnosis of group B strep during labour</p> <p>Methods of predictions and prevention of pre-eclampsia</p>	





# Appendix 7

## Questionnaire used in survey

University  
of Southampton

### **WESSEX INSTITUTE for HEALTH RESEARCH & DEVELOPMENT**

**An assessment of the impact of the NHS HTA Programme**

#### **Questionnaire**

Ethics No: 05/MRE02/11

Office use

ID:

Project Reference No:

Project Title:

Lead Researcher:

Questionnaire completed by: .....

Position (if not lead investigator) .....

Date: .....

## A. USE OF THE RESEARCH IN THE RESEARCH SYSTEM

- A1.** Has participation in this research led to additional formal qualifications for any members of the project team or is it likely to do so?

Yes No 

Qualification	Year		Contribution from specific project (please tick)		
	Gained	Expected	Considerable	Moderate	Small

- A2.** Have the project findings or methodology or theoretical developments generated subsequent research **by members of the team**?

Yes No 

- A2b.** If so, please give details of further grants, if any, and describe the contribution of your original project to securing these funds.

Funder	Amount	The importance of the project to securing later funding		
		Considerable	Moderate	Small

- A3.** If you are aware of any significant ways in which your HTA project has contributed to further research conducted **by others**, please indicate.

Research Team	Project title/topic	The importance of your project to the further research		
		Considerable	Moderate	Small

- A4.** Please describe any contribution to further research that you have listed in **A2** or **A3** that is of particular importance:

## B. USE OF RESEARCH FINDINGS IN HEALTH SYSTEM POLICY/DECISION-MAKING

Please note, questions about applications of the findings by practitioners are covered in the next section.

- B1.** Research findings can be used in policy making at any level (e.g. national, local trust or unit, professional, administrative or managerial) of the health service.

Have the findings from your project already been used in any such ways?

Yes

No

- B2.** Are there any reasons for expecting the findings to be used for future policy/decision-making?

Yes

No

- B3.** If you have replied **Yes** to either **B1** or **B2** please give details below of the use and/or expected use including: the level at which policies/decisions were influenced; the importance of the project's findings to the adoption of the policy(ies); and any supporting evidence<sup>1</sup> – please attach documents where relevant or give references to them.

## C. APPLICATION OF THE PROJECT FINDINGS THROUGH CHANGED BEHAVIOUR

- C1.** Have the findings from your project already led to changes, either directly or through the application of research-informed policies, in the behaviour of practitioners, managers etc, or in the involvement of health service users or the wider public?

Yes

No

- C2.** Do you expect the findings to influence practitioner or managerial behaviour or involvement of health service users or the public in the future?

Yes

No

- C3.** If you replied **yes** to either **C1** or **C2** please specify: **a)** the level at which any change occurred (e.g. local-institution, local-network, national); **b)** how important the research findings were in changing behaviour; and **c)** any evidence (such as surveys of practitioners) to support claims that such changes in behaviour were caused by the research findings – please attach documents where relevant or give references to them.

A – Level of change	
B – Importance in changing behaviour	
C – Evidence of change	

<sup>1</sup> Evidence of the policy relevance could take many forms including: statements by policy makers; NICE guidance; citing of the findings in a clinical guideline from a national or local professional group or in a National Service Framework; inclusion of the findings in a contract or in a document from an audit, an inspectorial or an evaluative body; the establishment of a working group to examine the implications or implementation of the findings etc.

## D. FACTORS INFLUENCING THE UTILISATION OF RESEARCH, INCLUDING DISSEMINATION

**D1.** Please state approximately how many conference/workshop presentations have been made based on the research findings to:

Primarily academic audiences: .....

Primarily practitioner audiences: .....

Primarily service user audiences: .....

**D2(a)** Were any of these presentations, or any other dissemination activities, particularly important in achieving utilisation of the project's findings?

Yes

No

Not/applicable

**D2(b)** If yes, please describe:

**D3(a)** Were any aspects of interaction with potential users particularly important?

Yes

No

Not/applicable

**D3(b)** If yes, please describe:

**D4(a)** Has the research been utilised?

Yes

No

Not/applicable

**D4(b)** Describe any other factors that account for the research being utilised, or for the lack of utilisation. These could include the timeliness or quality of the research, the research findings being taken up by the key stakeholders etc.

**D5.** How far was utilisation of the project enhanced by association with the NHS HTA programme?

Not at all

A little

Moderately

Considerably

Extensively

## E. PUBLICATIONS

Please check the accompanying publications list from information held on the NCCHTA database and:

**E1.** Please delete any that were not, at least partially, a result of funding from the HTA programme for the specific project identified on the previous page;

**E2.** Please list any additional publications from this HTA project on the form below;

**E3.** Please tick, for each publication on either list, either the left hand side of the entry or, for those added, tick the box if the publication contained an acknowledgement of HTA programme funding;

**E4.** Please use one of the following letters to categorise each publication – either by the side or using the appropriate box:

- A = peer-reviewed journal article
- B = journal editorial
- C = journal letter
- D = published abstract
- E = book
- F = chapter
- G = non-peer reviewed article
- H = published conference proceedings
- I = publicly available full report (including HTA monograph)
- J = newspaper (please specify)
- K = radio
- L = television
- M = other (please specify)

Additional Publications	Category A-M	HTA Fund Acknowl.	
		Yes	No

## F. COMMENTS

**F. Project Ref No:**

If you wish to make any further comments about your project or this questionnaire please use the space provided below.

Thank you very much for your assistance.

Please return the completed questionnaire in the return addressed **FREEPOST** envelope provided to: D. Coulson, HTA Impact Project, Wessex Institute for Health Research and Development, Mailpoint 728, Boldrewood, University of Southampton, Bassett Crescent East, Southampton SO16 7PX.

**ID No:**

# Appendix 8

## Sixteen case studies

The 16 case studies are listed in *Box 6*.

### Case study 1: A randomised controlled comparison of alternative strategies in stroke care (HTA study 93/03/26)

#### Stage 0: Needs assessment

Stroke was seen as the single most expensive disorder managed in general hospitals and the

burden was thought likely to increase. There were debates about how it should best be managed, with growing criticisms of the poor organisation of rehabilitation and a variety of initiatives being tried to provide more effective and organised stroke care. In the early 1990s, the Chief Investigator, Lalit Kalra, began studying ways of providing a more organised service. Studies were beginning to show that an organised stroke service resulted in better outcomes than conventional care such as that provided in a general medical ward, but there were questions as to exactly what form

#### BOX 6 The 16 projects selected for case studies

##### Primary research

*Case study 1:* A randomised controlled comparison of alternative strategies in stroke care (HTA study 93/03/26)

*Case study 2:* Effectiveness of counselling, cognitive-behaviour therapy and GP care for depression in general practice (HTA study 93/07/66)

*Case study 3:* Randomised evaluation of alternative electrosurgical modalities to treat bladder outflow obstruction in men with benign prostatic hyperplasia (BPH) (HTA study 94/04/09)

*Case study 4:* Efficacy and cost-effectiveness of physiotherapy for children less than four years old with cerebral palsy (HTA study 94/42/06)

*Case study 5:* Randomised controlled multiple treatment comparison to provide a cost-effectiveness rationale for the selection of antimicrobial therapy in acne (HTA study 94/48/03)

*Case study 6:* 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 (HTA study 95/07/19)

*Case study 7:* Psychological treatment in the regulation of long-term hypnotic drug use (HTA study 95/30/02)

*Case study 8:* Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain (HTA study 96/40/07)

*Case study 9:* Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography (HTA study 98/16/04)

##### Secondary research

*Case study 10:* Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis (HTA study 95/29/04)

*Case study 11:* Managing the dyspeptic patient: a systematic review and modelling exercise (HTA study 96/37/01)

*Case study 12:* Systematic review and evaluation of methods of assessing urinary incontinence (HTA study 99/29/02)

*Case study 13:* A systematic review to examine the impact of psycho-educational interventions on health outcomes and costs in adults and children with asthma (HTA study 01/16/02)

##### NICE Technology Assessment Reports (TARs)

*Case study 14:* The clinical effectiveness and cost-effectiveness of riluzole for the treatment of motor neurone disease (NICE TAR 00/01/01)

*Case study 15:* 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 (NICE TAR 00/13/01)

*Case study 16:* Effectiveness and cost-effectiveness of imatinib for first-line treatment of chronic myeloid leukaemia in chronic phase: a systematic review and economic analysis (NICE TAR 02/18/01)

this organised care should take. There was also pressure to manage a greater proportion of patients at home, but concerns about the feasibility, acceptability and costs of so doing. Kalra and colleagues developed a proposal to study the alternative ways of providing organised stroke care as part of the NHS R&D Programme on Cardiovascular Disease and Stroke. This was not funded, but the HTA Programme then invited tenders to study approaches to the effectiveness and resource implications of alternative organisational models of stroke rehabilitation (Priority area 93/3).

### Interface A: Project specification and selection

The team made a few additions to their existing proposal and submitted it to the HTA Programme, which accepted it. The objectives of the clinical evaluation included comparing a range of outcomes at 3, 6 and 12 months between three approaches to managing patients with non-severe stroke through organised stroke care.

Organisational aspects were to be described and the acceptability of various strategies evaluated. The aims of the economic evaluation were to collect data on service use (all agencies), accommodation and caregiver support in order to calculate the associated costs with each of the three modes of stroke rehabilitation and analyse the links between costs and outcomes.

#### Stage 1: Inputs

The study received £499,778 from the HTA Programme. This provided funding for a doctor, a physiotherapist and an occupational therapist and the specialist health economic and statistical input. The input of some further staff was provided by the Health Trust and the local Social Services department. The project Steering Group contained various stakeholders from the services involved and provided important support.

#### Stage 2: Processes

The study was conducted between October 1995 and March 1998. It was a prospective, single-blind RCT undertaken in patients recruited from a community-based stroke register. The health and social care needs of the district were provided for by a co-terminus hospital trust, a community health trust, a family health services authority and social services. Patients with severe strokes were excluded.

The three arms of the study were: *the stroke unit* that provided 24-hour care provided by a specialist multidisciplinary team based on clear

guidelines for acute care, prevention of complications, rehabilitation and secondary prevention; *the stroke team* that involved management on general wards with specialist team support to provide stroke assessments and advice to ward-based nursing and therapy staff on acute care, secondary prevention and rehabilitation aspects; and *domiciliary care* that consisted of management at home under the supervision of a GP and stroke specialist with support from a specialist team and community services for a maximum of 3 months.

Of the 979 patients entered on the stroke register during the period, 457 (47%) were randomised. Of these, 152 patients were allocated to the stroke unit, 152 to the stroke team and 153 patients to domiciliary stroke care (average age 76 years, 48% women). The groups were well matched for baseline characteristics, stroke type and severity, level of impairment and initial disability.

The primary measure was death or institutionalisation at 1 year. Secondary measures involved dependence, functional abilities, mood, quality of life, resource use, length of hospital stay and patient, carer and professional satisfaction.

### Stage 3: Primary outputs

#### Knowledge production

The main findings reported in the HTA were that, "Fifty-one (34%) patients in the domiciliary group were admitted to hospital after randomisation. Mortality and institutionalisation at 1 year were lower on stroke unit compared with the stroke team [21/152 (14%) versus 45/149 (30%),  $p < 0.001$ ] or domiciliary care [21/152 (14%) versus 34/144 (24%),  $p = 0.03$ ]. Significantly fewer patients on the stroke unit died compared with those managed by the stroke team [13/152 (9%) versus 34/149 (23%),  $p = 0.001$ ]. The proportion of patients alive without severe disability at 1 year was also significantly higher on the stroke unit compared with the stroke team [129/152 (85%) versus 99/149 (66%),  $p < 0.001$ ] or domiciliary care [129/152 (85%) versus 102/144 (71%),  $p = 0.002$ ]. These differences were present at 3 and 6 months after stroke.

"The total costs of stroke per patient over the 12-month period were £11,450 for the stroke unit, £9527 for the stroke team and £6840 for home care. More than half the total costs were incurred in the first 3 months. However, the mean costs per day alive for the stroke unit were significantly less than those for the specialist stroke team (£37.98 versus £50.90,  $p = 0.046$ ) patients, but no



different to those for domiciliary care patients. Costs for the domiciliary group were significantly less than for those managed by the specialist stroke team on general wards.” (Kalra and colleagues, 2005, pp. ix–x).

The authors concluded that, “Management of stroke patients on general medical wards, even with specialist team support, cannot be recommended because of the high mortality and dependence rate. ... a role for specialist domiciliary services for acute stroke was not supported ... the stroke unit is a more cost-effective intervention than either the stroke team or home care.” (Kalra and colleagues, 2005, p. x).

The monograph was not published until 5 years after the main article from the project, so it is the four articles in the *Lancet* and *Stroke* that are of most importance, although the second clinical article in the *Lancet* and the clinical article in *Stroke* were not included on the NCCHTA database. Kalra was also named as one of the collaborators, although not on the writing committee, for the Cochrane review in this field – see below. The authors thought it was important to acknowledge HTA funding on the articles because it enhanced the status of the work.

Evans A, Harraf F, Donaldson N, Kalra L. Randomised controlled study of stroke unit care versus stroke team care in different stroke subtypes. *Stroke* 2002;**33**: 449–55. [2004 impact factor, 5.7; at scoring, 16 cites; current, 18].

Evans A, Perez I, Melbourne A, Steadman J, Donaldson N, Kalra L. Can differences in processes explain different outcomes between stroke unit and stroke-team care? *Lancet* 2001;**358**:1586–92. [2004 journal impact factor, 21.7; at scoring, 34 cites; current, 43].

Evans A, Perez I, Melbourne A, Steadman J, Kalra L. Alternative strategies for stroke care: a randomised controlled trial of three strategies of stroke management and rehabilitation. *Cerebrovasc Dis* 2000;**10**:60. [no journal impact factor; at scoring, 1 cite; current, 1].

Kalra L, Evans A, Perez I, Knapp M, Donaldson N, Swift C. Alternative strategies for stroke care: a prospective randomised controlled trial. *Lancet* 2000;**356**:894–9. [2004 journal impact factor, 21.7; at scoring, 53 cites; current, 65].

Kalra L, Evans A, Perez I, Knapp M, Swift C, Donaldson N. A randomised controlled comparison of alternative strategies in stroke care. *Health Technol Assess* 2005;**9**(18). [at original scoring in 2005, 61 despatched, 2475 hits, no cites; current, 118 despatched, 14,055 hits, 3 cites].

Patel A, Knapp M, Evans A, Perez I, Kalra L. Alternative strategies for stroke care – cost-effectiveness and cost

utility analysis from prospective randomised controlled trial. *Stroke* 2004;**35**:196–203. [2004 journal impact factor, 5.7; current cites, 10].

The importance of an article is sometimes indicated by an accompanying editorial. The first *Lancet* paper was accompanied by an editorial that suggested the study should help influence policy, but that various guidelines had already described the importance of providing acute hospital care for people who had just suffered a stroke (Hacke, 2000). An editorial accompanying the cost-effectiveness study described its importance (Holloway and Dick, 2004).

#### **Research targeting, capacity building and absorption**

Kalra conducted two further studies, including one as part of the NHS R&D Programme on the Primary/Secondary Care Interface that drew moderately on the HTA Programme study and was funded to £260,000. Analysis of some of the citations to the original article indicate the study has informed studies by other teams.

One of the team received their MD based considerably on the work and another an MSc. It also helped the career development of several team members.

#### **Interface B: Dissemination**

About 10 presentations were given to academic audiences and others to professional and service user groups. They were important at a local and professional level as ways of presenting the detail. The papers are cited in several systematic reviews, including some Cochrane Reviews, especially the one by the Stroke Unit Trialists’ Collaboration *Organised inpatient (stroke unit) care for stroke* (SUTC, 2001). In this review, the study was one of only five studies for which outcome data were available for a comparison of different forms of organised stroke unit care and it was the only one that compared a dedicated stroke ward over a mobile stroke team. It was given the top grade for its methods. The authors of the review concluded: “Acute stroke patients should be offered organised inpatient (stroke unit) care which is typically provided by a coordinated multidisciplinary team operating within a discrete stroke ward” (SUTC, 2001).

#### **Stage 4: Secondary outputs: informing policy**

The study seems to have a considerable impact on policy at various levels – local, professional, national and international. *The national clinical*

*guidelines for stroke* from the Royal College of Physicians (RCP) cite both the Kalra 2000 paper directly and the Cochrane review (RCP, 2004). The guidelines also identify the evidence from the Kalra study as being the strongest category and stress that the recommendation that patients are admitted under the care of a specialist team for their acute care and rehabilitation should be the highest priority. The *National service framework (NSF) for older people* also recommends the provision of care by a specialist coordinated stroke team (Department of Health, 2001). It does not directly reference Kalra and colleagues (2000) but the study is thought to have influenced the framework and the Cochrane Review is cited. Similarly, a National Audit Office report (NAO, 2005) cites the Cochrane Review and the RCP Guidelines.

The Scottish Intercollegiate Guidelines Network (SIGN) Report 64, *Management of patients with stroke* (SIGN, 2002), recommends that acute stroke patients admitted to hospital should be treated in a multidisciplinary stroke unit, and cites the Cochrane Review (SUTC, 2001). It goes on to consider the option of arranging care in the patient's own home but recommends that: "Stroke patients who are dependent in activities of daily living should receive hospital-based care in organised stroke units." For this the main evidence given is Kalra and colleagues (2000).

Guidelines in several countries also cite papers from this study. The Stroke Council of the American Heart Association recently endorsed the guidelines from the Veteran Affairs/Department of Defense and these cite both Lancet articles as important evidence on the organisation of Stroke care (Duncan and colleagues 2005). For example, two sources of evidence are used to support the recommendation for 'organized and coordinated post-acute inpatient rehabilitation care'. One of these is Evans and colleagues (2001) and the two sources are given the top rating for quality of evidence.

### Stage 5: Applications by practitioners and public

The recommendations from the NSF carry considerable weight and are widely implemented, with audit procedures to check compliance. A key issue that arises is the counter-factual: how far would the changes in policy and practice have occurred without this study? In relation to the move towards stroke units, it provided strong evidence for the way policy and practice were beginning to go. What might be less clear is how

far the study slowed moves towards wider adoption of care at home, but it seems likely to be happening with the increased emphasis being placed on stroke patients being admitted to hospital and receiving stroke unit care. The national audit conducted in 2004 on behalf of the Intercollegiate Stroke Working Party showed that the proportion of stroke patients treated in a stroke unit had already increased from 36% to 47%.

### Stage 6: Impacts or final outcomes

The study showed various gains, especially reduced mortality, from the provision of care in specialist stroke units. It is therefore reasonable to suggest that following the widespread adoption of stroke units there has been a health gain. There will also have been reduced morbidity and increased patient satisfaction from the move away from care on general wards and increased provision of specialist units.

### Comments

In the case study, various documents referred to in the questionnaire and/or the interview were identified and examined, as were documents such as systematic reviews and guidelines appearing in the ISI list of publications citing one of the main articles from the study. Analysis of the context and wording of the citations received by the publications helped to indicate the strength of the impact made by the study on the policy documents. Although the information from both the questionnaire and from the case study indicated a high level of impact on policy, the case study provided a more detailed picture. The case study clarified that the project had some impact on the NSF, although not directly cited, but it also identified the impact on SIGN and the international impact.

This case study provides a good example of where a well-funded independent trial produced strong publications that were regarded as high-quality evidence in several important policy statements. Almost certainly the study fed through into impact on practice and actual health gain, but this is much more difficult to measure.

### References

Department of Health. *National Service Framework for Older People*. London: Department of Health; 2001. URL: <http://www.dh.gov.uk/assetRoot/04/07/12/83/04071283.pdf>.

Duncan PW, Zorowitz R, Bates B, Choi JY, Glasberg JJ, Graham GD, *et al*. Management of adult rehabilitation care: a clinical practice guideline. *Stroke* 2005;**36**:e100–43.

Hacke W. A late step in the right direction of health care. *Lancet* 2000;**356**:869–70.

Holloway R, Dick AW. Editorial comment – stroke cost-effectiveness research: are acceptability curves acceptable? *Stroke* 2004;**35**:203–4.

Kalra L, Evans A, Perez I, Knapp M, Swift C, Donaldson N. A randomised controlled comparison of alternative strategies in stroke care. *Health Technol Assess* 2005;**9**(18).

National Audit Office (NAO). *Reducing brain damage: faster access to better stroke care*. HC 452 Session 2005–2006. London: NAO; 2005.

Royal College of Physicians (RCP). *National Clinical Guidelines for Stroke*. 2nd ed. London: Prepared by the Intercollegiate Stroke Working Party, RCP; 2004.

Royal College of Physicians (RCP). *National sentinel stroke audit* 2004. London: Prepared on behalf of the Intercollegiate Stroke Working Party, RCP; 2005.

Scottish Intercollegiate Guidelines Network (SIGN). *Management of Patients with Stroke*. National Clinical Guideline No. 64. Edinburgh: SIGN, Royal College of Physicians; 2002.

Stroke Unit Trialists' Collaboration (SUTC). Organised inpatient (stroke unit) care for stroke. *Cochrane Database Syst Rev* 2001 (3).

## Case study 2: Effectiveness of counselling, cognitive-behaviour therapy and GP care for depression in general practice (HTA study 93/07/66)

### Stage 0: Needs assessment

Disorders involving symptoms of anxiety and depression are prevalent in the UK. The Principal Investigator of this study, Michael King, and others shared a growing interest in the way that counselling was expanding in the early 1990s and yet there was not a clear evidence base about its effectiveness. They conducted a pilot randomised trial in which GPs did the randomisation (King and colleagues, 1994). That did not really work because it was difficult for GPs to assess patients' preferences and conduct the randomisation. So the team went on to conduct a trial using routine randomisation. This was eventually published in the *Lancet* (Friedli and colleagues, 1997) and showed no difference between brief non-directive psychotherapy and usual general practice care, but patient recruitment was slow. King and colleagues were considering whether a Patient Preference Trial might be viable when the HTA Programme call 93/07 requested proposals to study the

effectiveness of counselling for patients with depression.

### Interface A: Project specification and selection

In response to the HTA Programme call, the team proposed a study to determine both the clinical effectiveness and cost-effectiveness of usual GP care compared with two types of brief psychological therapy in the management of depression and also mixed anxiety and depression in the primary care setting. In addition to including non-directive counselling, they also thought it important to add an arm based on the increasingly popular approach based on cognitive-behaviour therapy (CBT) for which evidence of efficacy was accruing.

Various practitioners, including counsellors from the pilot study, provided advice for the design of the study. It was principally a pragmatic RCT, but was accompanied by two additional allocation methods allowing patient preference: the option of a specific choice of treatment (preference allocation) and the option to be randomised between the psychological therapies only. The benefits of such a trial design are that it helps with patient recruitment because some patients have strongly held preferences and also it is appealing to the clinical professionals. It is, however, more expensive as larger numbers are required and there are also methodological objections to the inclusion of non-randomised participants. The HTA programme was seen as a particularly important funder in a field such as this where there would be no industry funding available. Only minor changes were made following the peer-review process.

### Stage 1: Inputs

The HTA funding was £426,908. There was no additional service support funding from the NHS. The team that had conducted the pilot study from the Royal Free combined with other researchers including from the National Primary Care R&D Centre, Manchester, and the Centre for Health Economics, York.

### Stage 2: Processes

The study was conducted between February 1996 and November 1997 in 24 general practices in Greater Manchester and London. A total of 464 eligible patients, aged 18 years and over, were referred by 73 GPs and allocated to one of the psychological therapies or usual GP care for depressive symptoms. The interventions consisted of brief psychological therapy (12 sessions maximum) or usual GP care. Of the 464 patients

allocated to the three treatments, 197 were randomised between the three treatments, 137 chose a specific treatment and 130 were randomised between the psychological therapies only. The patients underwent follow-up assessments at 4 and 12 months.

In the HTA monograph, the authors reported that given the impact of patient preferences in the trial was minimal, it did not seem likely that the large additional cost associated with preference design in trials would be a cost-effective use of research resources (King and colleagues, 2000). This, however, was a conclusion about the methods used and was not a problem that had to be overcome during this project. Despite the patient preference approach, there were still recruitment problems in some of the centres and the trial recruitment period had to be lengthened by 3 months.

### Stage 3: Primary outputs

#### Knowledge production

The main findings reported in the HTA were that, “At 4 months, both psychological therapies had reduced depressive symptoms to a significantly greater extent than usual GP care. ... These differences did not generalise to other measures of outcome. There was no significant difference in outcome between the two psychological therapies when they were compared directly using all 260 patients randomised to a psychological therapy by either randomised allocation method. At 12 months, the patients in all three groups had improved to the same extent. The lack of a significant difference between the treatment groups at this point resulted from greater improvement of the patients in the GP care group between the 4- and 12-month follow-ups. ... Similar outcomes were found for patients who chose either psychological therapy. Again, there were no significant differences between the two groups at 4 or 12 months. Patients who chose counselling were more satisfied with treatment than those who chose cognitive-behaviour therapy at 12 months. There were no significant differences in Beck Depression Inventory scores at either outcome point between participants who were randomised and those who chose each psychological therapy. No differences in direct or indirect costs between the three treatments were observed at either 4 or 12 months.” (King and colleagues, 2000, pp. iii–iv).

The authors concluded that, “In the primary care setting, non-directive counselling and cognitive-behaviour therapy were both significantly more effective clinically than usual GP care in the short

term. However, there were no differences between these three treatments in either clinical outcomes or costs at the 12-month follow-up. Psychological therapy provided in primary care was found to be a cost-effective method of reducing depressive symptoms in the short term, but the comparative benefits were relatively circumscribed and did not endure over the long term. Compared with usual GP care, no differences in overall costs were observed. The additional costs associated with providing practice-based psychological therapy were recouped due to savings in visits to primary care, psychotropic medication and other specialist mental health treatments.” (King and colleagues, 2000, p. iv).

The main implications for healthcare drawn by the authors were that, “based on the study’s observed equivalence in the clinical and economic outcomes of usual GP care compared with on-site psychological therapies in primary care, the commissioners of psychological services would be justified in considering additional factors when determining service configuration. These factors could include patient satisfaction, the preference of practitioners and staff availability.” (King and colleagues, 2000, p. iv).

In addition to the HTA monograph, a series of articles were produced, starting with the paper by Ward and colleagues (1999) setting out some methodological issues. Then there were linked papers in the *BMJ* giving the main clinical effectiveness and cost-effectiveness findings (Bower and colleagues, 2000 and Wood and colleagues, 2000). The two *BMJ* papers, plus a further one in the *BMJ* in 2002 describing a subsequent trial, were the ones listed on the NCCHTA database. In the questionnaire for our impact project, the 2002 paper was deleted from the list but the 1999 paper and two others were added: an analysis of the impact on the trial of patient–practitioner agreement by Gabbay and colleagues (2003) and another *BMJ* paper by Bower and colleagues (2003) reporting on a meta-analysis of data on costs from several trials including this one.

The linked, and well-cited, *BMJ* papers were thought to be the main outputs and the clinical effectiveness paper was runner-up in the Royal College of General Practitioner’s annual award for the publication with most impact. This an award for research papers nominated by GPs as having most impact on the primary care community.

Bower P, Byford S, Barber J, Beecham J, Simpson S, Friedli K, *et al.* Meta-analysis of data on costs from trials

of counselling in primary care: using individual patient data to overcome sample size limitations in economic analyses. *BMJ* 2003;**326**:1247–50. [2004 journal impact factor, 7; at scoring, 5 cites; current, 8].

Bower P, Byford S, Sibbald B, Ward E, King M, Lloyd R, *et al.* Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy and usual general practitioner care for patients with depression. II: Cost effectiveness. *BMJ* 2000;**321**:1389–92. [2004 journal impact factor, 7; at scoring, 23 cites; current, 26].

Gabbay M, Shields P, Bower P, Sibbald B, King M, Ward E. Patient-practitioner agreement: does it matter? *Psychol Med* 2003;**33**:241–51. [2004 journal impact factor, 2.9; at scoring, 2 cites; current, 0].

King M, Sibbald B, Ward E, Bower P, Lloyd M, Gabbay M, *et al.* 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. *Health Technol Assess* 2000;**4**(19). [at scoring in 2005, 799 despatched, 15,364 hits, 18 cites; current, 801 despatched, 23,300 hits, 23 cites].

Ward E, King M, Lloyd M, Bower P, Friedli K. Conducting randomized trials in general practice: methodological and practical issues. *Br J Gen Pract* 1999;**49**:919–22. [2004 journal impact factor, 1.6; at scoring, 10 cites; now, 11].

Ward E, King M, Lloyd M, Bower P, Sibbald B, Farrelly S, *et al.* Randomised controlled trial of nondirective counselling, cognitive-behaviour therapy and usual general practitioner care for patients with depression. I: Clinical effectiveness. *BMJ* 2000;**321**:1383–8. [2004 journal impact factor, 7; at scoring, 57 cites; current, 70].

### **Research targeting, capacity building and absorption**

The study helped target future work of team members in various ways, including a trial of teaching GPs skills in CBT. The main recommendation in their HTA report in relation to methodological research concerned patient preferences and they observed that the HTA Programme had already commissioned a systematic review of the effects of participants' and professionals' preferences in RCTs. In fact, this HTA review was being conducted by King and colleagues and has now reported (King and colleagues, 2005). It cost £125,217. One member of the team, Peter Bower, has been the lead author on several subsequent Cochrane Reviews in this field, including most recently one in which the original HTA study, and one other study, were jointly given the highest quality ranking of the eight studies included (Bower and Rowland, 2006).

The primary study has generated considerable international interest and, therefore, probably helped to inform a range of further studies by other researchers.

No formal qualifications were gained based upon research undertaken for the project, but it was seen as a success and several of the project team went on to receive Chairs and this study was probably one of many factors involved. It also encouraged some of the counsellors involved in the study to become more interested in research.

### **Interface B: Dissemination**

About 10 presentations to academic audiences were given by the Principal Investigator; other members of the team made further presentations. Presentations included ones given at the Royal College of Psychiatry and at the conference of the British Association for Counselling and Psychotherapy (BACP), which practitioners and service users attend. King is also a member of the Research Committee of BACP. The work was cited in the relevant edition of *Clinical Evidence* (Butler and colleagues, 2005) and in several Cochrane Reviews.

### **Stage 4: Secondary outputs: informing policy**

The study has had some influence on policy having been cited in various policy documents from national and professional bodies. The *National service framework (NSF) for older people* cites Ward and colleagues (2000) as the sole reference to support the statement: "Counselling in primary care may also be effective for depression at the less severe end of the spectrum" (Department of Health, 2001, para. 7.27). The Department of Health's Evidence Based Clinical Practice Guideline, *Treatment choice in psychological therapies and counselling*, relies mainly on systematic reviews but it reports the findings of the study as the main supplementary evidence in the section showing that psychological treatment has been found effective in the treatment of depression (Department of Health, 2001a, para. 2.1.1) The study is also cited in the NICE Clinical Guideline 23: *Depression – management of depression in primary and secondary care* (NICE, 2004). Given the range of diagnoses used in the study (i.e. anxiety and depression), it was not included as one of the 18 studies used, for the guideline, in the review of CBT for patients with a primary diagnosis of depression. Nevertheless, it was included in the review of counselling and the guideline says of the paper by Ward and colleagues (2000): "this is an

important primary care-based study” (NICE, 2004, para. 6.2.3.2). The study has been used by bodies such as the BACP; for example, it was highlighted in two press releases from BACP that promoted the role of counselling. All the various examples of impact on policy documents would support the suggestion that the study has had an impact on commissioning so as to increase the availability of counselling.

### Stage 5: Applications by practitioners and public

There is anecdotal evidence that the use of counselling has been encouraged by this study, especially as the study showed it was no more expensive. The award by the RCGP is an indication that the study had an impact, and was viewed as being important, but it provides no information about the degree of impact. It is difficult to know exactly what would have happened without the study because, as stated above in the section ‘Needs assessment’ (p. 125), the trend was already towards greater use of counselling.

### Stage 6: Impacts or final outcomes

The study showed the greater benefit of counselling and cognitive-behaviour therapy in the short term and such benefits to any one individual could be fairly large. Therefore, not only could there be a considerable health gain because of the total numbers involved, but also there could be a benefit to the economy in terms of a significant number of days off work avoided. It is impossible to make any calculations of actual gains, however, unless the impact of the study on practice could be assessed in some way.

### Comments

This case illustrates that in some instances the questionnaire provides a more accurate list of publications associated with a specific study than does the NCCHTA database. In addition to the interview and review of the project outputs, it was useful to undertake desk analysis and locate and analyse the relevant documents. These included some, such as the NICE Guidelines, that had been referred to in the questionnaire and interview, and others that had not but which were identified through web searches. Again, it is difficult to assess the counter-factual and know how far this study made a difference to any increase in counselling, but the fact that it is the only study referenced in the relevant section of the NSF could be significant. This would strengthen the view that this study was important in providing evidence to support a direction in which an increasing number of practitioners were interested in going.

### References

- Bower P, Roland N. Effectiveness and cost effectiveness of counselling in primary care. *Cochrane Database Syst Rev* 2006(3).
- Butler R, Carney S, Cipriani A, Geddes J, Hatcher S, Price J, *et al.* Depressive disorders; cognitive therapy. *Clin Evid* 2005. URL: [http://www.clinicalevidence.com/ceweb/conditions/meh/1003/1003\\_I10.jsp](http://www.clinicalevidence.com/ceweb/conditions/meh/1003/1003_I10.jsp)
- Department of Health. *National service framework for older people*. London: Department of Health; 2001. URL: <http://www.dh.gov.uk/assetRoot/04/07/12/83/04071283.pdf>.
- Department of Health. *Treatment choice in psychological therapies and counselling*. Evidence Based Clinical Practice Guideline. London: Department of Health; 2001a.
- Frieli K, King MB, Lloyd M, Horder J. Randomised controlled assessment of non-directive psychotherapy versus routine general-practitioner care. *Lancet* 1997;**350**:1662–5.
- King M, Broster G, Lloyd M, Horder J. Controlled trials in the evaluation of counselling in general practice. *Br J Gen Pract* 1994;**44**:229–32.
- King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, *et al.* Conceptual framework and systematic review of the effects of participants’ and professionals’ preferences in randomized controlled trials. *Health Technol Assess* 2005;**9**(35).
- National Institute for Clinical Excellence. *Depression – management of depression in primary and secondary care*. NICE Clinical Guideline 23. London: NICE; 2004.

## Case study 3: Randomised evaluation of alternative electrosurgical modalities to treat bladder outflow obstruction in men with benign prostatic hyperplasia (BPH) (HTA study 94/04/09)

### Stage 0: Needs assessment

Benign prostatic hyperplasia (BPH) is a disease of elderly men. In 1990 it cost the NHS an estimated £60–£91 million, 0.4% of the total spend. Although the prevalence of BPH has been put at 26–57% of men in the relevant age group, less than 1% were referred to a urologist.

The most common surgical treatment for bladder outflow obstruction is transurethral resection of the prostate (TURP), which uses a thin wire resectoscope loop to heat and disrupt cells. Although the mortality associated with TURP is low (0.1–0.25%), short-term morbidity is put at

18%, including bleeding, fluid absorption and urinary tract infection. Longer term morbidity is much higher, with 50–70% suffering retrograde ejaculation, and 5–30% erectile dysfunction.

The HTA Programme call 94/04 identified the need for a comparison of new and existing treatments for BPH. There were various new alternatives including transurethral vaporisation of the prostate (TUVP), which was thought to have similar effects but with less blood loss.

### **Interface A: Project specification and selection**

Once the HTA Programme call was made, the Principal Investigator, Christopher Fowler, put together a proposal with colleagues including a trainee surgeon who became the research fellow. As only two previous RCTs comparing TURP and TUVP had been reported, one a single-centre trial and the other excluding patients with retention, the case was made for a multicentre pragmatic trial of these two approaches. It was suggested that many of the newer means of treating BPH required expensive new equipment and that, therefore, before investing in such new equipment it would be important to evaluate cheaper electrosurgical methods which may be similarly efficacious.

Fowler was chair of the Standing Committee on Instruments of British Association of Urology Surgeons, and tutor and director of the minimal access therapy training unit of the Royal College of Surgeons, and therefore well placed to propose such a study at that time.

### **Stage 1: Inputs**

The team received £300,000 from the HTA Programme. Fowler and the research fellow were based at St Bartholomew's hospital, London. The other surgeons were based at other hospitals. In addition to the research fellow, there was a full-time research nurse at St Bartholomew's and two more in the contributing centres. Up to 10 surgeons were involved. The research team did not think that this study would have been funded by commercial companies. Most studies into new variations of technologies are small and cannot be considered as full RCTs. In the review of the evidence about BPH in *Clinical Evidence*, it is stated that, "rapid changes in techniques and few controlled trials with adequate follow up make comparisons between TURP and newer surgical techniques difficult." (Webber, 2004). In this instance there was substantial funding but, nevertheless, the funding was still probably insufficient and led to reduced numbers during the course of the trial.

The study was set up in parallel with the Prostate Trials Office (PROTO) based at the Bristol Urological Unit, also funded by NCCHTA and charged with providing a project steering group and data handling services for the trial described here. PROTO received considerable funding but in the end support for this project was its main activity and PROTO is described in the monograph as having provided critical input at several stages.

### **Stage 2: Processes**

The multicentre RCT was conducted between March 1997 and August 1999 with recruitment from four centres in South East England. A total of 235 men were recruited, all of whom had previously been assessed as requiring surgery for lower urinary tract symptoms deemed to be due to BPH. Randomisation was performed by a data monitoring team at PROTO, stratified to ensure even distribution of symptomatic and retention patients in each arm of the trial.

In addition to a realisation that the resources were insufficient to allow the original target levels of recruitment, there were also problems in conducting the study in the NHS because of capacity problems and need to give priority to emergency cases.

### **Stage 3: Primary outputs**

#### **Knowledge production**

There was a delay of 10 years from commissioning to publication of the monograph. Reasons for the delay included links with PROTO which had to be inaugurated before the study could start, problems with recruitment and, later, delays in receiving comments on the submitted monograph and in the process of amending the monograph. The team were very unhappy about these delays, only some of which they thought were due to them.

In the monograph, the team reported that: "TURP and TUVP were both effective in producing a clinically important reduction in IPSS [International Prostate Symptom Score] and positive change in the IPSS QoL question. The success rate for relief of symptoms, defined as a >5 reduction in IPSS at 6 months, was 85% for TURP and 74% for TUVP. Neither the success of the treatment nor the change in aggregated IPSS was significantly different between the groups. The improvement was sustained to 24 months after treatment with no significant difference between the groups. The effectiveness of both treatments was also equivalent when assessed through improvement in objective measures of urinary

tract function, reduction in prostate size and the change in health questions of SF-36. There was no change from baseline for other domains of SF-36 or EuroQoL.

“... The absolute incidence of adverse events was similar between the two groups. The incidence of severe or prolonged bleeding was less with TUVP, as evidenced by the need for blood transfusion and the drop in haemoglobin level 24 hours postoperatively.

“... TURP and TUVP are broadly equivalent in direct NHS resource use. In particular, staff costs, theatre usage and capital equipment costs are the same.” (Fowler and colleagues, 2005, p. x).

Therefore, the main conclusions drawn by the authors were that TURP and TUVP are equivalently effective in improving the symptoms of benign prostatic enlargement. As a result, replacement of TURP by TUVP would not produce a significant cost benefit to the NHS unless a reduction hospital inpatient stay of at least 1 day could be secured.

There was only one journal article besides the monograph, partly because the Principal Investigator did not believe in multiple publications of the same data but also because the technology had moved on over the 10 years and the issue researched had become a non-question. It was, nevertheless, published in the journal viewed as most important by British urology surgeons to inform their clinical practice (Jones and colleagues, 2006).

Fowler C, McAllister W, Plail R, Karim O, Yang Q. Randomised evaluation of alternative electrosurgical modalities to treat bladder outflow obstruction in men with benign prostatic hyperplasia. *Health Technol Assess* 2005;9(4). [at scoring in 2005, 57 despatched, 1509 hits, no cites; current, 112 despatched, 6448 hits, 1 cite].

McAllister WJ, Karim O, Plail RO, Samra DR, Steggall MJ, Fowler CG. Transurethral electrovaporisation of the prostate: is it any better than conventional transurethral resection of the prostate? *Br J Urol Int* 2003;91:211–14. [2004 journal impact factor, 2.1; at scoring, 3 cites; current, 4 cites].

### **Research targeting, capacity building and absorption**

The team have made no other bids for research due to the Principal Investigator moving on to other type of work. This study may have made urologists less keen on research – partly due to one untoward event in the trial. Nevertheless, the research fellow wrote his Master of

Surgery dissertation on the study for his degree in 2005.

### **Interface B: Dissemination**

Several presentations were made to practitioner audiences, including to the British Association of Urological Surgeons. In addition to appearing in the journal most widely read by British urology surgeons, and viewed as important for informing clinical practice by over 90% of them, it was also cited in the review in *Clinical Evidence* as one of the studies supporting the conclusion that, “RCTs found no significant difference in symptom scores between transurethral resection and transurethral incision or between transurethral resection and electrical vaporisation.” (Webber, 2004, p. 1128).

### **Stage 4: Secondary outputs: informing policy**

The findings came too late to make much, if any, impact on policy.

### **Stage 5: Applications by practitioners and public**

The findings came too late to make much, if any, impact on practice.

### **Stage 6: Impacts or final outcomes**

There are unlikely to have been any final impacts.

### **Comments**

The interview provided a fuller account of issues raised in the questionnaire about the difficulty of conducting this type of study on small improvements to technologies, especially in areas where there are frequent advances, and the problems with the various delays. There are also issues to do with studies that do not produce positive findings. All these are issues that could be important for the HTA Programme. A web search undertaken as part of the case study also identified a review in *Clinical Evidence* that provided further evidence on some of these issues. Furthermore, the interview came after the research fellow had obtained his MS degree and therefore information was gathered on this.

### **References**

Fowler C, McAllister W, Plail R, Karim O, Yang Q. Randomised evaluation of alternative electrosurgical modalities to treat bladder outflow obstruction in men with benign prostatic hyperplasia. *Health Technol Assess* 2005;9(4).

Jones TH, Hanney S, Buxton MJ. The journals of importance to UK clinicians: a questionnaire survey of



surgeons. *BMC Med Inf and Decision Making* 2006;6:24.  
URL: <http://www.biomedcentral.com/content/pdf/1472-6947-6-24.pdf>

Webber R. Benign prostatic hyperplasia. *Clin Evid* 2004;11:1119–38.

## Case study 4: Efficacy and cost-effectiveness of physiotherapy for children less than four years old with cerebral palsy (HTA study 94/42/06)

### Stage 0: Needs assessment

The HTA Programme call 94/42 set out that young children with cerebral palsy should have their needs properly addressed.

### Interface A: Project specification and selection

The Principal Investigator, Michael Weindling, was a consultant neonatologist whose research interests were in the prevention and treatment of acquired brain damage in newborn babies. He had supervised the PhD of a physiotherapist who was keen to respond to the HTA Programme call, so he put together a research team. Their original proposal covered the benefits from additional physiotherapy support for young children with spastic cerebral palsy. At the peer-review stage they were asked to give greater emphasis to assessing the needs of families and their requirements for support. The final draft of the HTA report was unpublished at the time the case study was completed, but NCCHTA and the lead author have kindly allowed access to it. That draft explained that the eventual proposal was that the physical and family support functions should be separately looked at in a trial. The trial investigated whether in the short and medium term additional support from a physiotherapy assistant improved physical function in young children with spastic cerebral palsy, and additional support from a family support worker improved family functioning. In addition, the study examined the needs of the families and the factors affecting child and family functioning in relation to services received and outcomes.

In the light of subsequent difficulties with the acceptability of the methods used, it was suggested that possibly in a complex, non-pharmaceutical, area such as this, there should have been greater liaison between the commissioners and the researchers at the project specification stage to get agreement as to appropriate methods.

### Stage 1: Inputs

The project received £344,093 from the HTA Programme and considerable further support from the NHS, for example through the supervision provided by senior physiotherapists in the various centres and the supervision of family support workers that was provided by a clinical psychologist. In addition to Weindling and the physiotherapist who first encouraged the application, the main members of the research team included psychologists and a health economist. Over the period of the study, three physiotherapy assistants and two family support workers were employed along with coordinators of the study.

### Stage 2: Processes

Conducted between March 1997 and September 2002, this was a multi-centre RCT with blinded assessments and a cost-effectiveness analysis. The children studied had spastic cerebral palsy that was the consequence of perinatal adversity. All were less than 4 years old on entry to the study. Randomisation was to (a) a group who received extra physiotherapy from a physiotherapy assistant, (b) a group who received standard physiotherapy and (c) a group where the child received standard physiotherapy and the family was also visited by a family support worker. Seventy-six families completed the intervention period. Forty-three families were reassessed 6 months after the end of the intervention and 34 of these after a further 6-month period.

In conducting the study, quantitative and qualitative methods were used and, as is always possible, some time after the study had started an improved technique became available for assessing a child's disability. The child outcome measures used were:

- motor functioning (Gross Motor Function Measure)
- developmental status (Griffiths Mental Developmental Scales)
- adaptive functioning (Vineland Scales).

The family outcome measures were:

- self-reported maternal stress (Parent Stress Index)
- level of family needs
- parental satisfaction.

In the process of the project, it became clear that the families had considerable support from various sources, in addition to any family support

that the physiotherapists had traditionally provided along with their actual physiotherapy. Therefore, although the roles of the physiotherapist were separated out in the study and in one arm the physiotherapy assistants were providing just extra physiotherapy, this did not mean that the families were receiving no support. It would probably have been unethical to have had a control group receiving no interventions. These problems were compounded by the variability in the conditions of the children with cerebral palsy.

These various issues raise some doubts about the appropriateness of conducting a traditional RCT in these circumstances and contributed to the problems with delays at various stages of the study. There were also, however, some problems with recruitment and some of the families appear not to have been reassessed.

### **Stage 3: Primary outputs**

#### ***Knowledge production***

The main findings from the RCT reported in the draft HTA monograph were largely negative: “There was no evidence that additional physical therapy for 1 hour a week for 6 months by a physiotherapy assistant improved any child outcome measure in the short or medium term. Intervention by a family support worker did not have a clinically significant effect on parental stress or family needs.” (Weindling and colleagues, 2007). In terms of costs, over the 6-month period the total cost of services for each child ranged from £250 to £6750, with higher costs associated with children with more severe impairments.

In addition, there was multivariate analysis across the board and in this no significant relationship was found between measures of intensity of services received by the children and families and the main outcome measures. Low-functioning children, in terms of both motor and cognitive function, were more likely to receive more services in terms of range and frequency. The qualitative analysis indicated that parents generally reported high satisfaction ratings after all interventions and some stated that the interventions had benefited the child and/or the family. There was therefore a discrepancy between the perceptions of these parents and the objective, quantitative measurements. The family support workers identified a small number of families who were experiencing considerable family problems, but who had not been referred for appropriate

support by any other agency. (Weindling and colleagues, 2007).

Some of the main conclusions from the authors, therefore, were that the findings “provide support for the current literature that there was no evidence that additional intervention (in this case by a physiotherapy assistant or family support worker) helped the motor or general development of young children with spastic cerebral palsy. Nor was there any quantitative evidence that providing extra family support helped levels of parental stress and family needs. ... In addition, no significant association was found between the intensity of the local services provided and any outcome measure, other than a slight association with lowered family needs. ... A small group of families with complex family problems needed more service input.”

In terms of recommendations, the authors suggested that overall it appeared that more funds were not needed. However, families who might particularly benefit from family-focused intervention were currently not appropriately targeted. They also argued that, ‘If the physiotherapist is to be the key professional (keyworker) for a child with cerebral palsy, guidance should be given on how to explain to parents that a higher level of physical therapy may not necessarily be warranted. A key worker for the child with cerebral palsy and his family needs to understand what family support entails, their role in providing it and acting as gatekeepers for referral to other agencies. If this really is an important part of the paediatric physiotherapist’s role, appropriate training and resources need to be provided.’ (Weindling and colleagues, 2007).

For a variety of reasons, there were delays in submitting the HTA monograph and subsequently in the process of review: it was subject to various requests for revision related to factors including the acceptability of the qualitative data. The authors delayed submitting journal articles until after final acceptance of the HTA report. There have been two published abstracts based on presentations given at the annual scientific meeting of the Royal College of Paediatrics and Child Health.

#### ***Research targeting, capacity building and absorption***

There have been limited opportunities for the study to influence further work, although the

authors set out various recommendations for further work that reflected some of the issues identified. In the light of the variations in current provision of support and of needs, they recommended that research was needed to examine what 'sufficient' levels of provision of therapy might be for which children and which families. They also suggested that future studies again adopt a multifaceted approach which is likely to be more suitable than a simple RCT.

No research degrees were achieved as a result of the project, but several team members have worked together on research applications in other topics.

### Interface B: Dissemination

There have been the two presentations referred to above and three seminars to physiotherapists whilst the study was in progress. A more substantial event to disseminate the findings has also been delayed until the report was accepted.

### Stage 4: Secondary outputs: informing policy

It is too early for the work to have informed any policies.

### Stage 5: Applications by practitioners and public

As yet it is unlikely that there will have been much impact on practitioners.

### Stage 6: Impacts or final outcomes

Again, it is unlikely that there will have been any final outcomes, although there would be the potential for more effective use of resources if the study helps to prevent the provision of additional family support apart from for the families on which support should be targeted.

### Comments

The case study illustrates the problems of trying to make an assessment of impacts from a study where, for a variety of reasons, there have been considerable delays. The discussion about the causes of some of these delays could, especially in combination with findings from some other case studies, provide lessons for the HTA Programme.

### Reference

Weindling M, Cunningham C, Glenn S, Tudor Edwards R, Reeves D. Efficacy and cost-effectiveness of additional therapy for young children with spastic cerebral palsy. *Health Technol Assess* 2007;**11**(16).

## Case study 5: Randomised controlled multiple treatment comparison to provide a cost-effectiveness rationale for the selection of antimicrobial therapy in acne (HTA study 94/48/03)

### Stage 0: Needs assessment

Acne is one of the most common skin disorders in young people. Having acne can give rise to feelings of embarrassment, loss of self-esteem and depression, and also physical symptoms (such as soreness and pain) associated with individual lesions. Most people with acne are treated in primary care. GPs have at least 30 different acne preparations to choose from, which can be prescribed singly or in combination, yet there are virtually no good **comparative** data to guide them or their patients to make the best choice in terms of efficacy, cost-effectiveness, compliance, tolerability and overall patient satisfaction. Antibiotic resistance in the bacteria implicated in acne pathogenesis may be associated with a reduction in clinical efficacy, and some antibiotic preparations may be more likely to promote resistance than others. The HTA Programme call 94/48 therefore identified this as a priority for research.

### Interface A: project specification and selection

The HTA Programme call came at a time when Hywel Williams, who eventually became the Principal Investigator, was looking to expand his research on clinical trials in the field of dermatology and he formed a collaboration with world leaders in acne research from Leeds University, Bill Cunliffe and Anne Eady. The patient group, the Acne Support Group, made a useful contribution to the study design from an early stage. The proposed study consisted of an RCT to assess the relative efficacy and cost-effectiveness of 11 of the most commonly used antimicrobial preparations for treating mild to moderate facial acne in the community. In addition, it was proposed to study whether pre-existing bacterial resistance to the prescribed antibiotic resulted in reduced efficacy and whether some antimicrobial regimens were less likely to give rise to resistant propionibacterial strains. Some minor improvements were made following the peer-review process.

### Stage 1: Inputs

The project finally received funding of £585,000 from the HTA programme after an additional sum

was added and there was the usual support for clinicians from the NHS. The multi-disciplinary team had a range of experience including, in addition to the expertise in dermatological research and acne, experience in microbiology, primary care and health economics.

## Stage 2: Processes

Participant recruitment took place between July 1998 and April 2000 for this RCT that used parallel comparative groups and a pragmatic design with intention-to-treat analysis. Major recruitment difficulties and high dropout rates prompted an early decision in consultation with the HTA Programme to restrict the study to just five treatment groups, instead of the intended 11. The five were chosen so that several questions important for the NHS could be addressed about the effectiveness and cost-effectiveness of leading therapies.

Because matched placebos would have been prohibitively expensive to produce, blinding of study participants was only partially achieved. Assessors were blinded to the intervention status of participants. Primary care practices and colleges in and around Nottingham and Leeds and one practice in Stockton-on-Tees were involved.

Eventually 649 people aged 12–39 years were recruited, all of whom had mild to moderate inflammatory acne of the face. This consisted of between 127 and 132 per treatment group which matched the original target per group, but the dropout rate was higher than expected. All existing acne treatments (oral and topical) were stopped for 4 weeks before the study.

## Stage 3: Primary outputs

### Knowledge production

The HTA monograph reported that the treatment differences for the primary outcomes between the five groups were small: “The percentage of participants with at least moderate improvement was 53.8% for minocycline (the least effective) and 66.1% for the combined erythromycin/benzoyl peroxide formulation (the most effective) . . . . Similar efficacy rankings were obtained using lesion counts, acne severity scores and global rating by assessor. Benzoyl peroxide was the most cost-effective and minocycline the least cost-effective regimen for treating mild to moderate inflammatory acne of the face . . . . The efficacy of oxytetracycline was similar to that of minocycline, but at approximately one-seventh of the cost. For all regimens, the largest reductions in acne severity were recorded in the first 6 weeks (around

45–50% of participants with at least moderate improvement). . . . The clinical efficacy of both tetracyclines was compromised in participants colonised by tetracycline-resistant propionibacteria. None of the regimens promoted an overall increase in the prevalence of antibiotic-resistant strains. Systemic adverse events were more common with the two oral antibiotics. Local irritation was more common with the topical treatments, particularly benzoyl peroxide. Residual acne was present in most participants (95%) at the end of the study.” (Ozolins and colleagues, 2005, p. x).

The authors concluded, “The response of mild to moderate inflammatory acne to antimicrobial treatment in the community is not optimal . . . . Perhaps the single most important finding of this study is that the topical antimicrobial therapies performed at least as well as oral antibiotics in terms of clinical efficacy. Benzoyl peroxide was the most cost-effective and minocycline the least cost-effective therapy for facial acne . . . . In addition to causing fewer systemic adverse events, topical preparations are less likely to induce resistance in other common bacteria, a finding that may be important for reducing the more widespread problem of bacterial resistance in the community . . . . Even though benzoyl peroxide was the most cost-effective treatment, it was associated with a greater frequency and severity of local irritant reactions. The results suggest that the use of a combination of topical benzoyl peroxide and erythromycin gives rise to less irritation and better quality of life. There was little difference between erythromycin plus benzoyl peroxide administered separately and the combined proprietary formulation in terms of efficacy or local irritation, except that the former was nearly three times more cost-effective.” (Ozolins and colleagues, 2005, pp. x–xi).

Both main outputs, the HTA monograph and the article in the *Lancet*, took a long time to be published. In the case of the *Lancet* article, it involved successful appeal against initial rejection. Acknowledging HTA Programme funding in the *Lancet* article was viewed as important and a way of emphasising that the research was independent of any pharmaceutical funding. The rapid increase in the number of hits for the HTA report indicates a considerable level of interest in this report.

Ozolins M, Eady EA, Avery A, Cunliffe WJ, O’Neill C, Simpson NB, *et al.* Randomised controlled multiple treatment comparison to provide a cost-effectiveness rationale for the selection of antimicrobial therapy in

acne. *Health Technol Assess* 2005;**9**(1). [at scoring in 2005, 60 despatched, 6119 hits, 1 cite; current, 122 despatched, 17,268 hits, 2 cites].

Ozolins M, Eady EA, Avery AJ, Cunliffe WJ, Po ALW, O'Neill C, *et al.* Comparison of five antimicrobial regimens for treatment of mild to moderate inflammatory facial acne vulgaris in the community: randomised controlled trial. *Lancet* 2004;**364**:2188–95. [2004 journal impact factor, 21.7; at scoring, 3 cites; current, 10].

Haynes M, Dada J. Contribution of nurses' involvement in acne research. *Nurs Stand* 2002;**16**:33–6. [no cites].

### **Research targeting, capacity building and absorption**

With the experience gained, particularly in this HTA trial, an Acne Research Group has been formed by Williams and colleagues to take forward research on a range of issues. These include antibiotic resistance which was one of the issues identified in the research recommendations in the HTA report. The group is independent of funding from the pharmaceutical industry. Some of the research nurses on the project have gone on to gain promotion and bring a research awareness to their clinical posts.

### **Interface B: Dissemination**

Various presentations were given to academic, professional and patient groups. There was also an active approach towards the media and considerable coverage ensued, for example, in the national and local radio, the *Yorkshire Post*, and in magazines and other news outlets in the USA and Canada. This coverage was noted by the HTA Programme in its 2005 update.

The first Cochrane Review of acne therapies predated the publications from this study (Garner and colleagues, 2003), but it is not clear why the study was not included in the article on acne vulgaris therapies published in *Clinical Evidence* in November 2005 with a search date of June 2005 (Purdy, 2005). It is understood that further UK reviews of acne are in the pipeline and that these draw extensively on the HTA study.

### **Stage 4: Secondary outputs: informing policy**

One of the difficulties in the acne field is that there appears to be a lack of official bodies with policy-making responsibility. There has been no guideline on acne *per se* from NICE, despite the potential for the findings to lead to cost savings. In December 2001, however, NICE issued *Referral advice: a guide to appropriate referral from general to specialist services* (NICE, 2001) and this included

advice on acne along with 10 other conditions. It stated that a treatment should be reviewed after about 2 months. Although this NICE document did not include any references, and anyway predated the publications from this study, Williams was listed as a member of the advisory group for the section on acne and was able to confirm that the traditional textbook 'review at 6 months' guidance was shortened to around 2 months as it was clear even at that stage from the study that if a treatment was going to help someone with acne, most of the response would be noticeable early on. [Nevertheless, the advice remained to continue therapy for a period of several months if it was initially found to be effective.]

The PRODIGY Guidance on acne vulgaris published in July 2006 makes extensive reference to the HTA report, citing it on at least seven occasions. For example, this guidance for healthcare professionals states: “**Review acne after 6 weeks** if there is little risk of scarring during this period. Most of the treatment effect of anti-acne drugs occurs during this time period [Ozolins and colleagues, 2005].” (PRODIGY, 2006).

The *Lancet* paper has been cited by two papers in which authors reviewed the literature on acne treatment. One in the *New England Journal of Medicine* makes clear the importance of the HTA Programme study but also stated, “There are currently no formal up-to-date, evidence-based guidelines available” (James, 2005, p. 1470). There was, however, a paper containing recommendations prepared by the European Expert Group on Oral Antibiotics in Acne (Dreno and colleagues, 2004). It cited an early version of the HTA report on several occasions.

### **Stage 5: Applications by practitioners and public**

Williams changed his practice and now uses topical rather than oral antibiotics for mild to moderate facial acne and reviews patients after 6 weeks rather than 3 months. There is anecdotal evidence that some other clinicians are similarly changing their practice and, given the policy recommendations described above, it would be reasonable to assume that there could be an increasing change in practice in line with the findings from the study. The only evidence that has been identified in this case study relates to falling sales of Minocin. Here, the counter-factual has to be considered because it was suggested that sales were already falling as a result of an article in the *BMJ* claiming to have identified a syndrome associated with the drug. Furthermore, the

Cochrane Review found no reliable RCT evidence to justify the continued first-line use of minocycline, “especially given the price differential and the concerns that still remain about its safety” (Garner and colleagues, 2003, p. 2). A further reduction in the use of minocycline might be expected following an article in August 2006 in *Drugs and Therapeutics Bulletin* that quoted extensively from the HTA Programme study (DTB, 2006).

It is possible that the *Lancet* article has had some direct impact on some of the clinicians who read it. It is also possible that there has been some shift in attitude towards the treatment of acne as a result of this major trial having been funded by the NHS and reported in the *Lancet*: it might have helped increase the seriousness with which the condition is regarded in primary care. Nevertheless, there are great difficulties in attempting any measurement of the degree of application of the findings by practitioners.

### Stage 6: Impacts or final outcomes

As reported in the HTA monograph, acne has a prevalence reaching almost 100% among adolescents. Therefore, any improvements in health gain and/or cost-effectiveness resulting from the trial could potentially bring widespread benefits. The trial showed that some treatments were more cost-effective than others and to the extent that the trial has had some impact on practice there should have been a small gain. Larger benefits could arise if further policy guidelines are produced in this field and further research is conducted on some of the topics identified in the study. In particular, there could be very important benefits in the wider health system if a move towards topical treatments helped reduce antibiotic resistance, and also the incidence of internal side-effects such as thrush and gastrointestinal upsets inherent when taking systemic antibiotics.

### Comments

This case study built on information from the questionnaire and interview. Desk analysis was undertaken and papers citing the publications from the project were identified and those that were reviews were accessed, as was the relevant NICE guideline on referral behaviour. In addition to the reviews identified by searching ISI's Web of

Science, a search on *Google Scholar* identified the review by the European expert group. The dramatic increase in the number of hits for the HTA monograph between the time of scoring and the time the case study was written up illustrates the potential importance of considering the time elapsed between the date of publication of research findings and the assessment of the impact. The presence of a specific body, sometimes called a 'receptor body' (Hanney and colleagues, 2003), to receive and use research findings can be a very important determinant of the level of research use. In the acne field, the lack of such a body is claimed to be a limitation on the utilisation of the findings from this study. This case study also highlights the advantages of providing HTA funding in an area where many previous studies were not independent of pharmaceutical funding.

### References

- Dreno B, Bettoli V, Ochsendorf F, Layton A, Mobacken H, Degreef H. European recommendations on the use of oral antibiotics for acne. *Eur J Dermatol* 2004;**14**:391–9.
- Drug and Therapeutics Bulletin. Is minocycline overused in acne? *Drug Ther Bull* 2006;**44**(8):60–2.
- Garner SE, Eady EA, Popescu C, Newton J, Wan Po A. Minocycline for acne vulgaris: efficacy and safety. *Cochrane Database Syst Rev* 2003(1).
- Hanney SR, Gonzalez-Block MA, Buxton MJ, Kogan M. The utilisation of health research in policy-making: concepts, examples and methods of assessment. *Health Res Policy Syst* 2003;**1**:2. URL: <http://www.health-policy-systems.com/content/1/1/2>
- James WD. Acne. *N Engl J Med* 2005;**352**:1463–72.
- National Institute for Clinical Excellence. *Referral advice: a guide to appropriate referral from general to specialist services*. London: NICE; 2001.
- Ozolins M, Eady EA, Avery A, Cunliffe WJ, O'Neill C, Simpson NB, *et al*. Randomised controlled multiple treatment comparison to provide a cost-effectiveness rationale for the selection of antimicrobial therapy in acne. *Health Technol Assess* 2005;**9**(1).
- PRODIGY. Acne vulgaris. *Prodigy Guidance*. Issued July 2006. London: NHS National Library for Health; 2006. URL: [http://www.prodigy.nhs.uk/acne\\_vulgaris/view\\_whole\\_guidance](http://www.prodigy.nhs.uk/acne_vulgaris/view_whole_guidance)
- Purdy S. Acne vulgaris. *Clin Evid* 2005. URL: <http://www.clinicalevidence.com/ceweb/conditions/skd/1714/1714.jsp>

## **Case study 6: 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 (HTA study 95/07/19)**

### **Stage 0: Needs assessment**

Previous research from the two lead applicants (Ann Oakley and Ian Roberts) indicated that social support could influence maternal depression, which in turn could, if mothers were happier, lead to a reduction in child injury. This led to the idea of conducting an RCT to test the effectiveness of providing a programme of social support postnatally. An application was made to the MRC that was alpha graded, but not funded. Then the HTA Programme issued call 95/07 requesting a cost-effectiveness study of alternative approaches to supporting families of high-risk babies.

### **Interface A: Project specification and selection**

In response to the HTA Programme call, Oakley and Roberts proposed the Social Support and Family Health (SSFH) study as an RCT which would compare maternal and child health outcomes for women offered support interventions through a programme of visits from health visitors trained in supportive listening, support health visitors (SHVs), with those for control women receiving standard services only. It aimed to measure the impact and cost-effectiveness of this strategy for providing support to mothers in disadvantaged inner-city areas.

There was extensive user involvement in the specification of this proposal. The idea of providing postnatal support had originally been raised by the participants in Oakley's earlier study, from the Social Science Research Unit of the Institute of Education, London, of supportive home visiting in high-risk pregnancy. In designing the SSFH, the researchers consulted the National Childbirth Trust (NCT). In the HTA's peer-review process, it was suggested that a further intervention arm be added that would involve the provision of postnatal support by lay individuals, which could potentially reduce the cost burden to the NHS and would meet a policy agenda related to the government's Sure Start initiative. In devising a suitable intervention, the team again consulted the groups such as the NCT. The Community Group Support (CGS)

intervention entailed the mother being assigned to one of eight community groups. The groups offered drop-in sessions, home visiting and/or telephone support. They made their standard package of services available to study women for 1 year.

There was also input into the study protocol from the local Health Authority, Camden and Islington, that provided funding for the health visitor intervention.

### **Stage 1: Inputs**

The eventual HTA grant was for £573,042, and some of the academics provided additional time. The amount from the HTA also included additional funding for a higher than expected need for interpreters. They were needed because an innovative part of the project was to aim to be as inclusive as possible in the recruitment of participants. Furthermore, the funding provided by Camden and Islington Health Authority for the health visitors' input was considerable. The lead applicants not only had wide experience in this field, but had shown a long-term commitment to this particular project. They recruited an appropriate team of health service researchers and health economists.

### **Stage 2: Processes**

The study ran from September 1998 to December 2001. Women living in deprived enumeration districts in the London boroughs of Camden and Islington were eligible for the trial if they gave birth between 1 January and 30 September 1999. In total, 731 women were recruited into the SSFH study, 58% of the 1263 eligible women. Half (364) were randomised into the control group and the remaining half allocated equally into the two interventions. Of the 184 allocated to the CGS intervention, 165 received offers of support but only 35 (19%) used CGS support. Of the 183 allocated to SHV intervention, 180 received an offer of support and 172 (94%) received at least one visit.

The authors describe the low uptake of the CGS intervention as a major limitation because the outcome data about effects of the two interventions were analysed on an intention-to-treat basis. The late request from funders for the policy issue of the non-professional support to be included as an extra arm meant that, "the best practice for designing trials, which involves exploratory work to develop interventions, was not possible for the CGS arm. Pilot work was not carried out; had it been, then it is likely that the low uptake of the CGS intervention would have

been detected, and this intervention would either have been abandoned or significantly modified.” (Wiggins and colleagues, 2004, p. 96). The problems created have had important implications, as described later.

Outcome data were collected through questionnaires distributed at 12 and 18 months. Here it is interesting that the response rate was very high across the arms of the trial: control, 90% after 12 months and 82% after 18 months; CGS, 89 and 85%; SHV, 90 and 80%. Process data were also collected. There was an integral economic evaluation.

### Stage 3: Primary outputs

#### Knowledge production

The main finding reported was that, “at both points there were no differences that could not be attributed to chance on the primary outcomes of maternal depression, child injury or maternal smoking. At both follow-ups there were differences in secondary outcomes: at the first follow-up, there was reduced use of general practitioners (GPs) by SHV children, but increased use of NHS health visitors and social workers by mothers; at the second follow-up, both CGS and SHV mothers had less use of midwifery services (fewer were pregnant), and SHV mothers were less worried about their child’s health and development . . . . Satisfaction with the intervention among women in the SHV group was high. Based on the assumptions and conditions of the costing methods, the economic evaluation found no net economic cost or benefit of choosing either of the two interventions.” (Wiggins and colleagues, 2004, p. ix).

In addition to the HTA monograph, there have been a series of publications aimed at different audiences, but the change in the design created difficulties for potential publication in the *BMJ* because concerns were expressed about the power of the eventual study. The various publications listed include one written solely by two of the health visitors who do not appear as authors on the main publications. The original list on the NCCHTA database included one entry that the respondent to the questionnaire had deleted as having been incorrectly added at NCCHTA. Several of the later publications were added on the questionnaire and had not been on the NCCHTA database. Three of the publications are in *Community Practitioner*, which is a professional journal but it has introduced a peer-review process. The publications are mostly too recent to have received many citations, but the large

number of hits already received by the HTA report indicates considerable interest in the findings.

Austerberry H, Wiggins M, Turner H, Oakley A. Evaluating social support and health visiting. *Community Practitioner* 2004;**77**:460–4. [no cites].

Oakley A, Strange V, Toroyan T, Wiggins M, Roberts I, Stephenson J. Using random allocation to evaluate social interventions: three recent UK examples. *Ann Am Acad Political Soc Sci* 2003;**589**:170–89. [at scoring, no cites; current, 1].

Oakley A, Wiggins M, Turner H, Rajan L, Roberts, Barker M. Including culturally diverse samples in health research: a case study of an urban trial of social support. *Ethnicity Health* 2003;**8**:29–39. [2004 journal impact factor, 0.7; at scoring, 1 cite; current, 1 cite].

Sawtell M, Jones C. Time to listen: an account of the role of ‘support’ health visitors. *Community Practitioner* 2002;**75**:461–63. [at scoring, no cites, current 2].

Turner H, Wiggins M, Oakley A, Austerberry H. Use of postnatal community services: RCT evidence. *Community Practitioner* 2005;**78**:11–15. [no cites].

Wiggins M, Oakley A, Roberts I, Turner H, Rajan L, Austerberry H, *et al.* 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. *Health Technol Assess* 2004;**8**(32). [at scoring in 2005, 67 despatched, 4366 hits, 2 cites; current, 131 despatched, 16,754 hits, 3 cites].

Wiggins M, Oakley A, Roberts I, Turner H, Rajan L, Austerberry H, *et al.* Postnatal support for mothers living in disadvantaged inner city areas: a randomised controlled trial. *J Epidemiol Community Health* 2005;**59**:288–95. [2004 journal impact factor, 2.8; no cites].

#### Research targeting, capacity building and absorption

The project helped to inform a subsequent study by Oakley and colleagues: *Teenage parenthood and social exclusion: a multi-method study* (Wiggins and colleagues, 2005). This received £175,000 from the Department of Health and included secondary analysis of some data collected in the SSFH study and further data collection from some of the participants. The main researcher has advised various subsequent projects, including being on the Advisory Group for an RCT from another team on peer-led infant feeding support.

The project helped advance the career of the main researcher and she is using the experience of the SSFH in her postgraduate teaching. The health visitors involved in the study have also become keen on advocating the use of research, and in some cases conducting it.



## Interface B: dissemination

The publications listed include ones aimed at academic audiences, but also there are ones, for example in *Community Practitioner*, targeted explicitly at practitioner audiences. The findings were also disseminated in a booklet for users and this was well received by the participants, other mothers and user groups. Furthermore, summaries of the report have been included on certain resource bases. Various presentations were given, including at an event for policy makers on postnatal care organised by the Royal College of Midwives and to parliamentarians on the relevant all-party group. Despite the activity reported here, most of the team were contract researchers and therefore did not have the time to disseminate the findings to the extent they would have wished.

The study is highly relevant for an updated Cochrane Review of RCTs of home-based support for socially disadvantaged mothers, but that had not been published by the summer of 2006. When it is eventually published this should not only help disseminate the findings of the SSFH study but also, as noted by Wiggins and colleagues (2005), set them in the context of the previous trials of the intervention.

## Stage 4: Secondary outputs: informing policy

One of the limitations on impact from the study is the lack of a single clear message; this also hinders attempts to identify any direct impact that might have arisen. Nevertheless, the paper in the *Journal of Epidemiology and Community Health* (Wiggins and colleagues, 2005) was cited in the draft NICE National Practice Guideline *Antenatal and postnatal mental health* that went out for consultation in July 2006 (NICE, 2006). Also, some of the findings from the SSFH study were used by the Royal College of Midwives.

The follow-on project on *Teenage parenthood and social exclusion* (Wiggins and colleagues, 2005) was part of a programme of research commissioned explicitly to inform the government's Teenage Pregnancy Strategy and it has fed into policy documents from the Government.

## Stage 5: Applications by practitioners and public

As noted, some of the articles, especially in *Community Practitioner*, were aimed at practitioners and there is considerable interest from health visitors when the findings are presented. Thus far, however, it seems difficult to go beyond such anecdotal evidence when considering how far the SSFH might have impacted on practice.

## Stage 6: Impacts or final outcomes

The study identified that the interventions could impact on individuals in different ways and showed that although the provision of the intervention was costly, it could also reduce the demand on GP services, thus producing off-setting cost savings. The limitations on identifying any impact made on practitioners by the SSFH study mean, however, that it is unrealistic to expect to be able to measure any final outcomes that might have resulted from the study in terms such as health benefits.

## Comments

This is a study to which the lead applicants were highly committed, based on the experience from previous research, and for which, although the MRC declined support, the HTA Programme call allowed a successful proposal to be submitted. The study highlights the difficulties of trying to meet precise policy agendas when conducting methodologically advanced research: adding an extra arm caused problems. Problems such as this are key issues facing the organisation of health research systems (Kogan and Henkel, 1983; Kogan and colleagues, 2006) and this case study should help to inform recommendations made to the HTA Programme by our impact assessment project. On the other hand, the additional funding provided by the HTA Programme for interpreters did facilitate the researchers' innovative strategy of attempting to be as inclusive as possible in the recruitment of participants.

The dramatic increase in the number of hits received by the HTA report illustrates the considerable interest in the findings. Furthermore, this increase during the course of the impact study, plus the draft NICE Guideline that appeared during the impact assessment, illustrates that it can be important to leave a reasonable time before attempting to assess research impact. In conducting the impact assessment, it was useful to undertake web-searches to supplement the information from the questionnaire and interview.

## References

- Kogan M, Henkel M. *Government and research: the Rothchild experiment in a government department*. London: Heinemann; 1983.
- Kogan M, Henkel M, Hanney S. *Government and research: thirty years of evolution*, 2nd ed. Dordrecht: Springer; 2006.
- National Institute for Health and Clinical Excellence. *Antenatal and postnatal mental health: consultation*. Draft National Practice Guideline. London: National Institute for Health and Clinical Excellence; 2006. URL: <http://www.nice.org.uk/page.aspx?o=340580>

Wiggins M, Oakley A, Roberts I, Turner H, Rajan L, Austerberry H, *et al.* 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. *Health Technol Assess* 2004;8(32).

Wiggins M, Oakley A, Sawtell M, Austerberry H, Clemens F, Elbourne D. Teenage parenthood and social exclusion: a multi-method study. London: Social Science Research Unit, Institute of Education and London School of Hygiene and Tropical Medicine; 2005.

## Case study 7: Psychological treatment in the regulation of long-term hypnotic drug use (HTA study 95/30/02)

### Stage 0: Needs assessment

There was widespread dissatisfaction with the cost and lack of effectiveness of the use of hypnotic drugs for the management of insomnia. Despite substantial developments in the science of psychological treatment of insomnia using CBT, there were limited advances in terms of how to apply it in primary care. The lead author of the HTA report, Kevin Morgan, constructed a multi-disciplinary consortium that applied to the MRC for a study to develop a costed service development model for application of a CBT package. The proposal was not funded. Then the HTA Programme invited tenders to study approaches to study the relative cost-effectiveness of alternative strategies for managing patients on long-term benzodiazepine medication who do not wish to cease medication (Priority area 95/30).

### Interface A: Project specification and selection

In response to the HTA Programme call, the team adapted their previous application and proposed a trial to assess: (1) whether psychological treatments for insomnia can be effectively delivered in routine NHS general practice settings by non-sleep specialists and (2) whether improvements in sleep quality achieved through psychological treatment can produce significant and sustained reductions in hypnotic drug use among long-term hypnotic users. The proposal received constructive feedback from the HTA Programme commissioning panel. In addition to making a few minor additions that were accepted, they also suggested that the project would be rather under-resourced and agreed to a request for further funding.

### Stage 1: Inputs

In addition to the £186,626 funding from the HTA Programme, about another 10% was added by the research team. The research team combined skills from the Sheffield Institute of General Practice and Primary Care, the Sheffield Health Economics Group and Nottinghamshire Healthcare Trust Department of Clinical Psychology.

### Stage 2: Processes

The patients were recruited between January 1999 and August 2001 and the 12-month assessments continued until November 2001. The study was designed as a pragmatic cluster RCT with two treatment arms (a CBT-treated 'sleep clinic' group, and a 'no additional treatment' control group), with post-treatment assessments starting at 3, 6 and 12 months. All patients entered the treatment receiving hypnotic drugs. The intervention consisted of six 50-minute sessions delivered by primary care counsellors. The main outcomes included global sleep quality, frequency of hypnotic drug use, mean dose of hypnotics consumed, health-related quality of life, NHS service costs and overall cost utility. The project over-ran because of recruiting problems, but it was important to have sufficient numbers. Of the 537 patients invited to join the trial, 209 (39%) agreed to do so.

### Stage 3: Primary outputs

#### Knowledge production

As probably the largest ever clinical trial in this field that had a health economic component alongside it, the knowledge production was significant. The key conclusions, as reported in the HTA monograph (Morgan and colleagues, 2004), were that: "At 3- and 6-month follow-ups, patients treated with CBT showed improved global PSQI scores ( $p < 0.002$  and  $p < 0.04$ , respectively), and improvements in the SF-36 dimensions of vitality at 3 months ( $p < 0.01$ ), and physical functioning ( $p < 0.04$ ) and mental health ( $p < 0.02$ ) at 6 months. CBT-treated patients also reported reductions in the frequency of hypnotic drug use (both  $p < 0.001$ ) compared with the control group, with many CBT-treated patients (29% at 3 months and 33% at 6 months) reporting zero drug use at the follow-up assessments. Clinical improvements were maintained within the CBT group at the 12-month follow-up, with PSQI scores ( $p < 0.01$ ) and the frequency of hypnotic drug use ( $p < 0.001$ ) continuing to show significant reductions relative to the control group ...

"The total cost of service provision was £154.40 per patient (1999/2000 prices). The mean

incremental cost per quality-adjusted life-year (QALY) at 6 months was £3418; this figure, within a range that has previously been considered to represent acceptable value for money by NHS decision-makers, was insensitive to changes in costs (varying from £3074 to £4679 per QALY when counsellor unit costs were changed). While the incremental gain in utility was not statistically significant, when combined with the incremental cost data, the probability that the cost per QALY of treatment would be considered cost-effective if decision-makers are willing to pay less than around £12,500 per QALY, is greater than 80%.” (Morgan and colleagues, 2004, pp. ix–x).

The authors therefore concluded that, “Despite chronic hypnotic drug use ostensibly to manage persistent insomnia, patients in the trial reported very high levels of sleep disturbance and very low levels of sleep quality. In routine general practice settings, psychological treatment for insomnia improves sleep quality, reduces hypnotic drug use, and improves health-related quality of life at a favourable cost among long-term hypnotic users with chronic sleep difficulties ... CBT for insomniacs should be considered by primary care commissioners and practitioners when implementing National Service Framework recommendations for benzodiazepine use”. (Morgan and colleagues, 2004, p. x).

In addition to the HTA report, there were three main peer-review articles, the third having been accepted just before the interview was conducted.

Dixon S, Morgan K, Mathers N, Thompson J, Tomeny M. Impact of cognitive behaviour therapy on health related quality of life among adult hypnotic users with chronic insomnia. *Behav Sleep Med* 2006;4:71–84. [no cites].

Morgan K, Dixon S, Mathers N, Thompson J, Tomeny M. Psychological treatment for insomnia in the regulation of long-term hypnotic drug use. *Health Technol Assess* 2004;8(8). [at initial scoring of projects in 2005, 65 despatched, 11,671 hits, 9 cites; current, 124 despatched, 20,888 hits; 10 cites].

Morgan K, Dixon S, Mathers N, Thompson J, Tomeny M. Psychological treatment for insomnia in the regulation of long-term hypnotic drug use: a pragmatic randomised controlled trial. *Br J Gen Pract* 2003;53:923–8. [2004 journal impact factor, 1.6; at scoring, 3 cites; current, 6].

Morgan K, Thompson J, Tomeny M, Dixon S, Mathers N. Predicting longer-term outcomes following psychological treatment for hypnotic-dependent chronic insomnia. *J Psychosom Res* 2003;54:21–9. [2004 journal impact factor, 2.8; at scoring, 4 cites; current, 5].

### **Research targeting, capacity building and absorption**

The HTA Programme project helped to target further work by the research team and by others. An essentially developmental project of £45,000 was funded by NHS University and the National Institute for Mental Health for England (NIMHE) as a direct follow-on from the HTA project. It entailed training people to become counsellors who could deliver the CBT. One of the counsellors on the research team is the principal tutor on the course. As a development project linked to service provision, the role of this project will be analysed later. The project, especially the economic analysis, also helped inform an RCT of nurse-administered CBT conducted by a team led by Colin Espie at Glasgow. Since publication of the findings, the work has generated considerable international interest and in 2005 helped inform the development of ‘Recommendations for a standard research assessment of insomnia’ (Buysse and colleagues, in press). In July 2006, two members of the original HTA team (Morgan and Tomeny) collaborated in a further related research development under the Cross-Council ‘New dynamics of ageing’ research programme. Using experience, training programmes and materials from the original HTA programme funded project, a major new trial of self-help CBT for insomnia symptoms among patients with chronic diseases will commence in 2007 (at a cost of £300,000).

No further research degrees were obtained directly from participation in this work, but the research assistant did go on to undertake a PhD within the Sheffield Institute of General Practice and Primary Care.

### **Interface B: dissemination**

In the survey response, the project reported 10 presentations to primarily academic audiences, six to practitioners and two to service users. Several important presentations have been made to international audiences, with particular interest coming from the USA, where there is more attention given to sleep medicine. Talking to local service providers has been complicated by the frequent turnover of staff and reorganisations.

Various articles in leading journals cite the work as part of discussions or recommendations on how to treat insomnia. An editorial in the *BMJ* on treating insomnia stated, “the current front runner for non-pharmacological treatment is cognitive behavioural therapy” (Holbrook, 2004), and it gave the HTA monograph as one of four references. Similarly, a seminar on insomnia in the

*Lancet* devoted a reasonable space to describing the findings, despite listing 174 references overall (Sateia and Nowell, 2004). An article on chronic insomnia in the Clinical Practice section of the *New England Journal of Medicine* also briefly refers to the study as part of the evidence supporting the recommendation to use cognitive-behavioural therapy (Silber, 2005). Finally, the study is briefly mentioned in a recent review in *Clinical Evidence* (Dunne and Montgomery, 2006).

#### Stage 4: Secondary outputs: informing policy

One of the problems of moving towards a policy on CBT is that it is not entirely clear who has responsibility for management in this field. A NICE TAR (Dundar and colleagues, 2004) conducted a systematic review and economic evaluation on newer hypnotic drugs for the short-term management of insomnia, but it said of the Morgan HTA report, "Although of considerable importance in relation to the use of hypnotics outside their therapeutic licence, this paper was not a drug versus drug comparison and was outside the scope of this review" (p. 39). This NICE TAR informed NICE Guidance on the use of newer hypnotic drugs in the short-term management of insomnia, NICE Technology Appraisal 77. Morgan was one of the individuals selected from client expert and patient advocate nominations to participate in the Appraisal Committee discussions and provide evidence, but the NICE Guidance did not contain any references beyond the TAR (NICE, 2004).

The trial outcomes have driven a programme of training in psychological treatments for insomnia for primary care professionals within the East Midlands. The decisions to establish these courses and their curriculum could be seen as a policy development, especially as the aim of the work is to develop a national model of training and practice for roll-out across the NHS.

#### Stage 5: Applications by practitioners and public

Most of the application of the particular CBT approach developed through the trial is taking place in the East Midlands through the activity of Morgan and colleagues. The various articles discussed above that cite the publications from this study could potentially encourage some further take-up of the approach.

#### Stage 6: Impacts or final outcomes

The results of the clinical trial suggest that there would be health gains from the adoption of CBT

for the management of insomnia. Furthermore, it is possible to argue that in the context of daytime sleepiness and hypnotic drug consumption being associated with reduced work performance and absenteeism (Morgan and colleagues, 2004, p. 1), then the introduction of an effective treatment could lead to wider economic benefits in the form of a healthier work force. The numbers who have so far been treated are small and therefore these benefits are largely potential rather than actual benefits at present. Nevertheless, the potential benefits will be very large because there are many prescriptions for hypnotic drugs for insomnia being written each year: over 10 million in 2000 (Morgan and colleagues, 2004, p. 1).

#### Comments

In our impact assessment, the questionnaire was comprehensively completed and this led to various avenues that could be further explored in the case study, including in the interview and in the examination of various documents located from the Internet. Various citing articles were also identified through searches on ISI and *Google Scholar*. This case study also highlights issues about the appropriate timing for impact assessments: several potentially important developments were in the pipeline at the time of the interview.

#### References

- Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. *Sleep* 2006;**29**:1155–73.
- Dundar Y, Boland A, Strobl J, Dodd S, Haycock A, Bagust A. *et al.* Newer hypnotic drugs for the short-term management of insomnia: a systematic review and economic evaluation. *Health Technol Assess* 2004;**8**(24).
- Dunne D, Montgomery P. Insomnia in the elderly. *Clin Evid* 2006;**15**:1–2.
- Holbrook AM. Treating insomnia. *BMJ* 2004;**329**:1198–9.
- Morgan K, Dixon S, Mathers N, Thompson J, Tomeny M. Psychological treatment for insomnia in the regulation of long-term hypnotic drug use. *Health Technol Assess* 2004;**8**(8).
- National Institute for Clinical Excellence. *Guidance on the use of zaleplon, zolpidem and zopiclone for the short-term management of insomnia*. Technology Appraisal Guidance No. 77. London: NICE; 2004.
- Sateia MJ, Nowell PD. Insomnia. *Lancet* 2004;**364**:1959–73.
- Silber MH. Chronic insomnia. *N Engl J Med* 2005;**353**:803–10.

## Case study 8: Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain (HTA study 96/40/07)

### Stage 0: Needs assessment

The Chief Investigator, Kate Thomas of the School of Health and Related Research (ScHARR) at the University of Sheffield, was interested in the potential of complementary therapies being offered on the NHS and identifying the type of evidence that would be necessary to move forward on this. She was already collaborating with Hugh MacPherson of the Traditional Chinese Medicine Foundation on a pilot study that included looking at back pain and acupuncture. This work involved working with practitioners rather than doing research 'on' them. The HTA Programme call 96/40 invited projects to analyse the effectiveness of acupuncture in the management of chronic pain in primary care. Part of the topic identification by the HTA was that as acupuncture is based on a wide variety of different theories and beliefs, it is unlikely that all forms would be equally effective.

### Interface A: Project specification and selection

In response to the HTA Programme call, a proposal was developed by a team including the researchers from ScHARR and the acupuncturist/researcher from York along with input from local GPs. They developed a proposed trial with the primary objective to test the hypothesis that a population of patients with persistent non-specific low back pain gained more long-term relief from pain when offered access to traditional acupuncture care than those offered conventional care only, and did so for equal or less cost. The team proposed a pragmatic effectiveness trial rather than an efficacy trial. This would entail the acupuncturists treating patients as they would in everyday practice and meant they would be treating the whole patient and not be working to a strict protocol. The commissioning panel asked the team to examine whether, as a result of that, there might be an acupuncturist effect. Therefore, instead of the proposed 200 patients with 100 in each arm, the proposal was changed to 240, randomised 2:1 in favour of the intervention arm in the hope that there would be at least 20 patients per acupuncturist. This request from the commissioning panel reflected some of the thinking in the topic identification.

### Stage 1: Inputs

In addition to the £221,129 from the HTA Programme, there was the usual input of academic time. Furthermore, the York Health Authority paid for the treatment costs for the acupuncture, which is not normally covered by the NHS. The research team included health service researchers, health economists and statisticians, the Traditional Chinese Medicine Foundation and a GP adviser. The advisory group provided useful input, including that from the patient representative.

### Stage 2: Processes

The pragmatic, parallel design RCT recruited patients prospectively, eventually over an 18-month period from April 1999 to January 2001. GPs used broad study entry criteria to identify eligible patients during consultations. The researcher applied the criteria more fully and eligible and interested patients received a home visit from the researcher to discuss the project in more detail. Patients in the experimental arm were offered the option of referral to individual acupuncture care in addition to having access to normal management for their back pain, at the discretion of their own GP. Full patient recruitment was achieved after a 6-month extension to the original 12-month period. Thirty-nine GPs from the York area participated in the trial. In total 241 patients were recruited into the study: 160 were randomised to the offer of acupuncture and 81 to usual GP management. In addition to baseline measures undertaken in the presence of the study researcher, postal questionnaires were sent after 3, 12 and 24 months. A study into consumer involvement in complementary research included this project in its analysis. It reported positively on how the patient representative had been able to contribute, for example by bringing about improvements in the questionnaire. (Patterson, 2003; URL: [http://www.hsra.ac.uk/Current\\_research/MRC\\_STFs/Final\\_consumers\\_in\\_CAM\\_research\\_report.pdf](http://www.hsra.ac.uk/Current_research/MRC_STFs/Final_consumers_in_CAM_research_report.pdf))

### Stage 3: Primary outputs

#### Knowledge production

The main results reported in the HTA monograph were that, "Analysis of covariance, adjusting for baseline score, found an intervention effect of 5.6 points on the SF-36 Pain dimension in favour of the acupuncture group at 12 months, and 8 points at 24 months. No evidence of heterogeneity of effect was found for the different acupuncturists. Patients receiving acupuncture care did not report any serious or life-threatening events. No treatment effect was found for any of the SF-36 dimensions other than Pain, or for the

ODI [Oswestry Disability Index]. Patients receiving acupuncture care reported a significantly greater reduction in worry about their back pain at 12 and 24 months compared with the usual care group. At 24 months, the acupuncture care group was significantly more likely to report 12 months pain free and less likely to report the use of medication for pain relief. The acupuncture service was found to be cost-effective at 24 months; the estimated cost per QALY was £4241 using the SF-6D scoring algorithm based on responses to the SF-36, and £3598 using the EQ-5D health status instrument. The NHS costs were greater in the acupuncture care group than in the usual care group. However, the additional resource use was less than the costs of the acupuncture treatment itself, suggesting that some usual care resource use was offset.” (Thomas and colleagues, 2005, p. ix). Based on the findings, the authors reported that “commissioners of musculoskeletal services would be justified in considering making GP referral to a short course of traditional acupuncture care available for a typical population of primary care attendees with persistent non-specific low back pain.” (Thomas and colleagues, 2005, p. x). There was, however, an issue of whether the results should be interpreted as being positive or negative and this led to a delay in the publication decision about a paper submitted to the *BMJ*.

The HTA report had not been published at the start of our impact assessment when the initial scoring of the project was conducted. The report was published in August 2005 and, as indicated by the number of hits received on the HTA website, it has attracted considerable attention. Two peer-reviewed articles had been published, but at time of interview the main papers were still under review. They have recently been published as linked papers in the *BMJ* and were accompanied by an editorial that referred to the study as a “well conducted” trial and highlighted its importance (Wonderling, 2006).

MacPherson H, Thorpe L, Thomas K, Campbell M. Acupuncture for low back pain: traditional diagnosis and treatment of 148 patients in a clinical trial. *Complement Ther Med* 2004;**12**:38–44. [2004 journal impact factor, 1.3; at scoring, no cites; current, 2 cites].

Ratcliffe J, Thomas KJ, MacPherson H, Brazier J. A randomised controlled trial of acupuncture care for persistent low back pain: cost effectiveness analysis. *BMJ* 2006;**333**:626–9. [2004 journal impact factor, 7; just published]

Thomas KJ, MacPherson H, Ratcliffe J, Thorpe L, Brazier J, Campbell M, *et al.* Longer term clinical and cost effectiveness of traditional acupuncture for

persistent low back pain in primary care. *Health Technol Assess* 2005;**9**(32). [at initial scoring in 2005, not published; current, 141 despatched, 18,262 hits, 3 cites].

Thomas KJ, Fitter MJ, Brazier JE, MacPherson H, Campbell M, Nicholl JP, *et al.* Longer-term clinical and economic benefits of offering acupuncture to patients with chronic low back pain assessed as suitable for primary care management. *Complement Ther Med* 1999;**7**:91–100. [no journal impact factor in 1999; at scoring, 8 cites; current, 9 cites].

Thomas KJ, MacPherson H, Thorpe L, Brazier J, Fitter M, Campbell MJ, *et al.* Randomised controlled trial of a short course of traditional acupuncture compared with usual care for persistent non-specific low back pain. *BMJ* 2006;**333**:623–6. [2004 journal impact factor, 7; just published].

Furthermore, a large number of the poster and oral presentations made about the project were published in some form, and this is further discussed below.

### **Research targeting, capacity building and absorption**

Further research by team members has been strongly informed by the original study in several ways. With funding from their respective universities, plus £8500 being contributed by *Backcare*, Thomas and MacPherson are conducting a 5-year follow-up of the patients in the original trial. Based partly on the HTA Programme trial, plus a wider range of research, MacPherson received a postdoctoral fellowship award from the National Coordinating Centre for Research Capacity and Development. The design of a trial at the University of Tromso was also informed by the HTA project.

In terms of capacity building, no research degrees were based on the work, but it has probably helped the career development of several of the team.

### **Interface B: dissemination**

There has been considerable effort to present the findings to academic and practitioner audiences (both in the back pain and acupuncture communities). This has included more than a dozen published oral or poster presentations at national and international conferences and meetings. It was planned that the main attempt at media coverage would surround the publication in the *BMJ*. This was successful; an article in the National Library for Health’s *Hitting the Headlines* series described the coverage in three national newspapers following the on-line publication (Centre for Reviews and Dissemination, 2006).

#### Stage 4: Secondary outputs: informing policy

Despite the considerable interest in the findings, it is too early for any impact on policy.

#### Stage 5: Applications by practitioners and public

There was some benefit to the acupuncturists involved in the study because they usually treat patients on a one-off basis with a variety of conditions, whereas in the trial they were forced to consider a series of patients with a common condition. Outside the trial, however, despite the considerable interest in the findings, it is too early for any impact on practice.

#### Stage 6: Impacts or final outcomes

The trial results suggest there could be some health gain from the wider adoption of this therapy. Back pain is recognised as a cause of lost production and adoption of the findings could possibly lead to some gains in the health of the workforce.

#### Comments

The case highlights the issue of timing of impact assessments because when the data about each project were taken from the database of the NCCHTA, the HTA monograph for this project had not been published. It was published during the course of the impact assessment and has been accessed a considerable number of times in its first year. Similarly, the linked papers in the *BMJ* were published at the very end of the process of conducting this case study. It was, nevertheless, too early to assess how the undoubted interest shown in the project might have translated into impact on policy or practice.

The considerable efforts that went into disseminating the findings from this study have been recorded above. This study was included, in the exploratory analysis undertaken for our impact assessment, on a list of those projects from the whole HTA Programme that have produced the highest number of outputs. The fact that, in this case, so many outputs – at least at the time the exploratory analysis was conducted – were presentations rather than articles possibly highlights the desirability of careful use of information from the NCCHTA database.

The acknowledgements in the HTA monograph recognise the advice of someone who played a key role in a previous trial of exercise-based management of low back pain. That trial was analysed in study of the impact from the research

funded by the ARC (Wooding and colleagues, 2005). Some of the detailed methods in the HTA study might, therefore, to some degree have been informed by aspects of the ARC-funded study. This highlights a point often made by scientists about the frequently cumulative nature of research advances. It also illustrates how there could be a growing body of knowledge about how to assess payback in relation to streams of research.

#### References

Centre for Reviews and Dissemination. Acupuncture for low back pain. Hitting the Headlines 18 September 2006. National Library for Health. URL: <http://www.library.nhs.uk/rss/newsAndRssArticle.aspx?url=http%3a%2f%2fwww.library.nhs.uk%2fresources%2f%3fid%3d177801>

Thomas KJ, MacPherson H, Ratcliffe J, Thorpe L, Brazier J, Campbell M, *et al.* Longer term clinical and cost effectiveness of traditional acupuncture for persistent low back pain in primary care. *Health Technol Assess* 2005;**9**(32).

Wonderling D. Acupuncture in mainstream care (Editorial). *BMJ* 2006;**333**:611–12.

Wooding S, Hanney S, Buxton M, Grant J. Payback arising from research funding: evaluation of the Arthritis Research Campaign. *Rheumatology* 2005;**44**:1145–56.

#### Case study 9: Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography (HTA study 98/16/04)

##### Stage 0: Needs assessment

Breast cancer is the most common cancer in women and approximately 6000 cancers are detected each year through the UK NHS Breast Screening Programme (NHSBSP). The programme, however, faced both increased demand and staffing problems. The use of computer-aided detection (CAD) for breast cancer screening was a proposed way of addressing this. The HTA Programme call 98/16 identified the need to assess the cost-effectiveness and appropriateness of using computer aids in the NHSBSP. The call stressed that for many questions posed by the HTA Programme the primary research required was an RCT, but it recognised that this might not be appropriate in this case and invited applicants to justify their choice if they proposed an alternative approach.

## Interface A: Project specification and selection

In response to the HTA Programme call, Paul Taylor and colleagues developed a proposal to determine the value of CAD for breast cancer screening through assessing the impact of the R2 ImageChecker® on the sensitivity and specificity of radiologists and film-reading radiographers in two experiments, referred to in the HTA report as study 1 and study 2 (Taylor and colleagues, 2005). The team took the view that the available evidence about CAD could not justify the setting up of an RCT for various reasons, including the need to get consent from 60,000 women and the delays while women returned for subsequent screening. Therefore, it was proposed that two sets of mammograms with known outcomes should be used. Participants in both studies would read the films with and without the benefit of the computer aid. The resulting data would be used in an economic evaluation. Some additional items were added as a result of the review process, but the methods remained the same.

There was a substantial delay in the commissioning process, however, because financial constraints meant that not all approved projects could be funded and the project was subject to a second review process.

### Stage 1: Inputs

The project received £278,000 from the HTA Programme, but this included additional funding of about £100,000 for an extension. In addition to Taylor, the Principal Investigator, the team included specialists in cancer, mammography, health economics and statistics.

### Stage 2: Processes

Study 1 was conducted between January 2001 and July 2002 and study 2 between September 2002 and April 2003. The studies were conducted at five screening centres: South-West London, Norfolk and Norwich, Luton and Dunstable, Worthing and Bristol. Thirty radiologists, five breast clinicians and 15 radiographers participated. All cases in the trial were digitised and analysed using the R2 ImageChecker version 2.2. Participants all received training on the use of CAD. In the intervention condition, participants interpreted cases with a prompt sheet on which regions of potential abnormality were indicated. The first set of 180 films, used in study 1, included 20 false-negative interval cancers and 40 screen-detected cancers. The second set of 120 films, used in study 2, was designed to be favourable to CAD: all 44 cancer cases had previously been

missed by a film reader and cancers prompted by CAD were preferentially included.

In terms of appropriateness of the research processes, perhaps had the CAD system been as effective as was claimed then the methods adopted would have enabled an effect to have been found. A further concern was that the whole process of applying to the HTA Programme and having to go through a further round, then conducting the detailed research and then the delays in the reviewing process of the HTA monograph meant that the HTA report did not appear until more than 6 years after the original call. In their HTA report, the authors note that such delay “is clearly inappropriate in a field where technology changes suddenly and rapidly” (Taylor and colleagues, 2005, p. 40). They went on to suggest that the existing HTA Programme funding mechanisms seemed inappropriate for the type of short, focused studies that answer specific questions about specific systems.

### Stage 3: Primary outputs

#### Knowledge production

The main findings in the HTA report were that, “No significant difference was found for readers’ sensitivity or specificity between the prompted and unprompted conditions in study 1 [... sensitivity with and without CAD is 0.76 to 0.80, for specificity it is 0.81 to 0.86 without CAD and 0.81 to 0.87 with CAD]. No statistically significant difference was found between the sensitivity and specificity of the different groups of film reader [... unprompted sensitivity of radiologists was 0.75 to 0.81, for radiographers it was 0.71 to 0.81, prompted sensitivity was 0.76 to 0.81 for radiologists and 0.69 to 0.79 for radiographers]. Thirty-five readers participated in study 2. Sensitivity was improved in the prompted condition (0.81 from 0.78) but the difference was slightly below the threshold for statistical significance (95% CI for the difference  $-0.003$  to  $0.064$ ). Specificity also improved (0.87 from 0.86); again, the difference was not significant at  $0.05$  (95% CI  $-0.003$  to  $0.034$ ). A cost-effectiveness analysis was performed based on data from studies 1 and 2. The analysis showed that computer prompting is cost-increasing.” (Taylor and colleagues, 2005, p. ix).

The team concluded that although the case for CAD as an element of the NHSBSP was not made, further evaluations of new CAD tools in routine use were under way and their results should be given careful attention. In particular, they



suggested that prompts may have an impact in routine use that is not detectable in an experimental setting.

There were three peer-reviewed articles in addition to the HTA monograph, but the third, which describes these two studies, and another, were published in 2005 and not recorded on either the NCCHTA database or the questionnaire and therefore were not included in the original scoring.

Taylor P, Champness J, Given-Wilson R, Johnston K, Potts H. Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography. *Health Technol Assess* 2005;**9**(6). [at initial project scoring in 2005, 57 despatched, 1710 hits, no cites; current, 113 despatched, 6087 hits, 2 cites].

Taylor P, Champness J, Given-Wilson R, Potts H, Johnston K. An evaluation of the impact of computer-based prompts on screen readers' interpretation of mammograms. *Br J Radiol* 2004;**77**:21–7. [2004 journal impact factor, 1.2; at scoring, 7 cites; current, 11].

Taylor P, Champness J, Given-Wilson R, Potts H, Johnston K. Assessing the impact of CAD on the sensitivity and specificity of film readers. *Clin Radiol* 2004;**59**:1099–105. [2004 journal impact factor, 1.5; at scoring, 1 cite; current, 3].

Taylor P, Given-Wilson R. Evaluation of computer-aided detection (CAD) devices. *Br J Radiol* 2005; **78** (special issue 1):S26–30. [2004 journal impact factor, 1.2; current, 3 cites].

The delays referred to above probably meant that by the time the monograph was published it was too late to make much impact. It might be significant that the article published earlier in the *British Journal of Radiology* has received some citations.

#### **Research targeting, capacity building and absorption**

The team received about £10,000 from the NHSBSP to complete some significant work related to the original project. The team also made the data from the study available to a group from City University, who undertook some further analysis on it and published the results (Alberdi and colleagues, 2004).

#### **Interface B: Dissemination**

Two presentations were made to academic audiences and nine to practitioner audiences; several presentations were published in some form. The team also talked to journalists and to the people running the Breast Screening Programme.

#### **Stage 4: Secondary outputs: informing policy**

As a study with negative findings, it is particularly complex to determine its exact impact on policy. Nevertheless, there is some evidence to support the view that the NHSBSP took some notice of the report because when it produced *Quality assurance guidelines for breast cancer screening radiology* (Liston and Wilson, 2005) the research was cited in a section at the end and a reference made to the findings. In the USA there has been much wider adoption of CAD and the fact that the NHSBSP has not yet recommended its introduction perhaps reflects some impact from the study. Meanwhile, negative results are being published by others, for example Hukkinen and colleagues (2006), who state that their negative findings are in line with those from the HTA Programme study.

#### **Stage 5: Applications by practitioners and public**

Meetings were held with practitioners during and at the end of the research and the articles were in fairly widely read journals, but the nature of the findings means that it is difficult to assess any impact on practitioners.

#### **Stage 6: Impacts or final outcomes**

Again there are limitations as to how far it is possible to consider any final outcomes. To the extent, however, that the study did help prevent the introduction of costly equipment that would probably not have produced much health gain, it could be argued that the study might have made some contribution towards efficiency in the NHS.

#### **Comments**

The authors' recommendation that "there should be a clearer and speedier route to commissioning evaluations of rapidly changing technologies" (Taylor and colleagues, 2005, p. x) reflects frustrations with the various delays described above. As part of the analysis of factors associated with the levels of impact achieved, this is clearly something that the HTA Programme should consider. In terms of maximising impact, it is important to try to ensure that studies are timely.

For the impact assessment itself, the interview provided various leads in terms of a further publication and an NHSBSP report that were subsequently identified in Internet searches. The interview also provided a steer on some key points, quotations about which could then be extracted from the monograph. The case study also perhaps illustrates the potential difficulties in achieving,

and/or identifying, impact from a study with negative findings.

## References

Alberdi E, Poyakalo A, Stringini L, Ayton P. Effects of incorrect computer-aided detection (CAD) output on human decision-making in mammography. *Acta Radiol* 2004;**11**:909–18.

Hukkinen K, Vehmas T, Pamilo M, Kivisaari L. Effect of computer-aided detection on mammographic performance: experimental study on readers with different levels of experience. *Acta Radiol* 2006;**47**: 257–63.

Liston J, Wilson R, (editors). *Quality Assurance Guidelines for breast cancer screening radiology*. NHSBSP Publication No. 59. London: NHSBSP; 2005.

Taylor P, Champness J, Given-Wilson R, Johnston K, Potts H. Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography. *Health Technol Assess* 2005;**9**(6).

## Case study 10: Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis (HTA study 95/29/04)

### Stage 0: Needs assessment

There is considerable mortality from coronary heart disease in people with familial hypercholesterolaemia (FH). In the majority of cases the disorder is caused by a mutation of the low-density lipoprotein receptor gene that impairs its proper function, resulting in very high levels of plasma cholesterol. Most people with FH are undiagnosed or only diagnosed after their first coronary event, but early detection and treatment with hydroxymethylglutaryl-coenzyme reductase inhibitors (statins) can reduce morbidity and mortality. The prevalence of FH in the UK population is estimated to be 1 in 500, which means that approximately 110,000 people are affected and yet at least 75% are undiagnosed. In call 95/27, the HTA Programme requested a systematic review to inform options for screening policy in this field.

### Interface A: Project specification and selection

The Principal Investigator, Margaret Thorogood, worked with colleagues including an expert in cardiovascular genetics, to put together a proposal in response to the HTA call. The peer-review

comments were positive on the proposed systematic review that aimed to: evaluate whether screening for FH was appropriate; determine which system of screening is most acceptable and cost-effective; assess the deleterious psychosocial effects of genetic and clinical screening for an asymptomatic treatable inherited condition; and assess whether the risks of screening outweigh potential benefits.

### Stage 1: Inputs

The team received £68,298 from the HTA Programme. Thorogood and the health economist and research fellow on the project were based at the London School of Hygiene and Tropical Medicine. The other applicants were based at the Centre for Cardiovascular Genetics, University College, London (Steve Humphries), the Department of Public Health and Primary Care, University of Oxford (Andrew Neil) and the Department of Social Medicine, University of Bristol (Helen Lambert).

### Stage 2: Processes

The review was conducted between January 1998 and June 1999. A systematic search of electronic databases was supplemented by handsearches including additional papers referenced in the search material identified. The inclusion criteria for the review covered studies of the mortality and morbidity associated with FH, the effectiveness and cost of treatment (ignoring pre-statin therapies in adults), and of the effectiveness or cost of possible screening strategies for FH. The search for papers on the psychological and social effects of screening for a treatable inherited condition was limited to the last 5 years because recent developments in genetic testing have changed the nature and implications of such screening tests. Papers focusing on genetic testing for FH and breast cancer were included. Papers relating to the risk of coronary heart disease with similarly modifiable outcome (non-FH) were also included.

A data assessment tool was designed to assess the quality and validity of the papers which reported primary data for the social and psychological effects of screening. Available guidelines for systematically reviewing papers concentrated on quantitative methods, and were of limited relevance. An algorithm was developed for use with both the qualitative and quantitative literature.

A model was constructed to investigate the relative cost and effectiveness of various forms of population screening (universal or opportunistic) and case-finding screening (screening relatives of known FH cases). All strategies involved a two-

stage process: first, identifying those people with cholesterol levels sufficiently elevated to be compatible with a diagnosis of FH, and then either making the diagnosis based on clinical signs and a family history of coronary disease or carrying out genetic tests. Cost-effectiveness was measured in terms of incremental cost per year of life gained.

### Stage 3: Primary outputs

#### Knowledge production

The authors report in the HTA monograph that it is appropriate to consider systematic screening for FH because diagnostic tests are reasonably reliable and acceptable, and treatment with statins substantially improves prognosis: "Case finding amongst relatives of FH cases was the most cost-effective strategy, and universal systematic screening the least cost-effective. However, when targeted at young people (16 year olds) universal screening was also cost-effective. Screening patients admitted to hospital with premature myocardial infarction was also relatively cost-effective. Screening is least cost-effective in men aged over 35 years, because the gains in life expectancy are small. The modelling results would support a combination of strategies. For example, universal systematic screening at 16 years could be carried out alongside both opportunistic screening of patients with an early myocardial infarction (men aged 16–34 years, women aged 16–54 years) and case finding for family members of index cases (men aged 16–34 years, women aged 16–54 years)." (Marks and colleagues, 2000, p. iv).

The report concluded that there was a lack of qualitative or quantitative evidence on the psychosocial effects of screening for FH and a lack of primary data (as opposed to modelling) on the effectiveness and cost implications of screening strategies. It therefore called for more research in these areas.

The HTA monograph was published relatively quickly in 2000 and the main publication in the *BMJ* followed in 2002. These timings might help explain why, unlike in most other case studies, the report has received more citations than the main journal publication. The *BMJ* article was included in the NCCHTA database and a later article in the *Journal of Public Health Medicine* was added in the questionnaire; this paper was based on funding from the NSC for modelling of the original data for which the HTA Programme is also acknowledged. During the course of the case study, a further publication was identified, in *Atherosclerosis*, which acknowledges the HTA

Programme as the main funding source. The citations to the publications indicate wide international interest in the findings, including, for example, an editorial in the *Medical Journal of Australia* that uses the work to support a call for a national programme in Australia to detect FH (Burnett, 2005).

Marks D, Wonderling D, Thorogood M, Lambert H, Humphries SE, Neil HAW. Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2000;**4**(29). [at initial project scoring in 2005, 338 despatched, 12,215 hits, 25 cites; current, 389 despatched, 21,287 hits, 34 cites].

Marks D, Wonderling D, Thorogood M, Lambert H, Humphries SE, Neil HAW. Cost effectiveness analysis of different approaches of screening for familial hypercholesterolaemia. *BMJ* 2002;**324**:1303–6. [2004 journal impact factor, 7; at scoring, 16 cites; now, 20].

Marks D, Thorogood M, Neil HAW, Wonderling D, Humphries SE. Comparing costs and benefits over a 10 year period of strategies for familial hypercholesterolaemia screening. *J Public Health Med* 2003;**25**:47–52. [2004 journal impact factor, 0.8; at scoring, 0 cites; now, 2].

Marks D, Thorogood M, Neil HAW, Humphries SE. A review on the diagnostic, natural history, and treatment of familial hypercholesterolaemia. *Atherosclerosis* 2003; **168**:1–14. [2004 journal impact factor, 3.8; now, 43 cites].

#### Research targeting, capacity building and absorption

The research was of considerable importance in helping to target several further projects from the team. They were funded with almost £30,000 from the NSC to conduct further modelling on the emerging data and over £23,000 from the British Heart Foundation. The work with the Foundation helped to identify research that was needed in this field. The team's further work in this field has led to further publications (Marks and colleagues, 2006).

The recommendations noted above for more research had a strong influence on the current pilot project on cascade screening that is described below as part of the policy response from the NSC.

The data contributed considerably to the research fellow's PhD and she was subsequently appointed to a joint post with a PCT and a university.

#### Interface B: Dissemination

About five presentations were made to academic audiences and five to practitioner audiences, but

other members of the team probably made further presentations, including to audiences in various countries. The team presented the findings to NSC when it was considering FH.

#### **Stage 4: Secondary outputs: informing policy**

The analysis of the project's impact on policy making is complex. It is cited in a range of national and international policy documents but its precise impact perhaps varies in different contexts. The *European guidelines on cardiovascular disease prevention in clinical practice* (Third Joint Task Force of European and Other Societies on Cardiovascular Disease Prevention in Clinical Practice, 2003) contain almost 800 references. In relation to the call for systematic screening of close relatives of individuals with FH, it cites a couple of studies and uses Marks and colleagues (2002) to support the statement: "Cost-benefits modelling based on data in the UK has demonstrated the effectiveness of cascade testing in the relatives of FH patients" (p. S37). Similarly, guidelines from the UK professional and patient bodies cite the study as demonstrating the effectiveness of cascade testing of the relatives of those FH (British Cardiac Society, British Hypertension Society, Diabetes UK, HEART UK, Primary Care Cardiovascular Society and The Stroke Association, 2005).

In terms of the screening policy of the NHS, the NSC policy position is to support the pilot project on cascade screening of the relatives of patients with confirmed FH that was announced in the genetics White Paper, *Our inheritance, our future* (Department of Health, 2003). How far the pilot project on cascade screening is based on the HTA report is not made clear in the White Paper because it contains no references, but the NSC lists the HTA report as the main evidence leading to its policy position (NSC, 2006). Furthermore, the Simon Broom FH Register states that the findings of the HTA study "were presented to the Department of Health and, in response, the White Paper on Genetics in the NHS published in 2003 committed funding to pilot a national cascade screening programme in five sites in England" (Department of Primary Health Care, University of Oxford; URL: [http://www.primarycare.ox.ac.uk/research/vascular/research/simon\\_broome](http://www.primarycare.ox.ac.uk/research/vascular/research/simon_broome)).

In information made available to the team assessing the impact of the HTA Programme, the report by Marks and colleagues (2000) was described as being the evidence behind the

recommendation that a pilot study of cascade screening should be introduced rather than full population screening. What is less clear in the impact assessment, however, is how far the recommendation not to introduce whole population screening might have been made even without the HTA Programme study. There was enthusiasm for more screening but probably not a real campaign for universal screening.

NICE guidelines are currently being developed for the identification and management of patients with FH and members of the team are involved in the process. In relation to identification of individuals with FH, the specific methods highlighted for investigation are opportunistic identification and cascade screening (NICE, 2006).

#### **Stage 5: Applications by practitioners and public**

The pilot studies are being introduced, but it is unclear whether the published guidelines are having any wider impact on practice.

#### **Stage 6: Impacts or final outcomes**

If the policy is successfully introduced there could be substantial health gain given that 75% of the estimated 110,000 people who have FH are undiagnosed and yet treatment is effective and delays or prevents the onset of coronary heart disease. Given the age of the patients, there would also be a considerable human capital benefit in terms of a healthier workforce.

#### **Comments**

This study has attracted considerable attention internationally in addition to the UK and informed recommendations from various professional bodies, including at the European level. Furthermore, it has had a direct impact on the policy of the NSC.

Some of the key information about this project's impact on the policymaking of the NSC was available to the impact assessment project team prior to undertaking the case study. This came from the list of HTA Programme projects that have been used by the NSC (see Appendix 6 of this report). Further details came from the questionnaire, the interview and various Internet searches that identified papers citing the main publications from the study. A review of these identified various clinical guidelines in which the study is cited.

#### **References**

British Cardiac Society, British Hypertension Society, Diabetes UK, HEART UK, Primary Care Cardiovascular

Society and The Stroke Association. JBS2: Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. *Heart* 2005;91 Suppl. V:v1-53.

Burnett JR. Familial hypercholesterolaemia: a look back, a look ahead. (Editorial). *MJA* 2005;182:552-3.

Department of Health. *Our inheritance, our future: realising the potential of genetics in the NHS*. Cm5791-II. London: Department of Health; 2003.

Department of Primary Health Care, University of Oxford. *Simon Broome FH register*. URL: [http://www.primarycare.ox.ac.uk/research/vascular/research/simon\\_broome](http://www.primarycare.ox.ac.uk/research/vascular/research/simon_broome).

Third Joint Task Force of European and Other Societies on Cardiovascular Disease Prevention in Clinical Practice. European Guidelines on Cardiovascular Disease Prevention in Clinical Practice. *Eur J Cardiovasc Prev Rehabil* 2003;10 Suppl. 1:S1-78.

Marks D, Wonderling D, Thorogood M, Lambert H, Humphries SE, Neil HAW. Screening for hypercholesterolaemia versus case finding for familial hypercholesterolaemia: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2000;4(29).

Marks D, Thorogood M, Neil SM, Humphries SE, Neil HAW. Cascade screening for familial hypercholesterolaemia: implications of a pilot study for national screening programmes. *J Med Screen* 2006;13:156-9.

National Institute for Health and Clinical Excellence. *Familial hypercholesterolaemia: draft scope*. London: NICE; 2006. URL: <http://www.nice.org.uk/page.aspx?o=340235>.

National Screening Committee. *Hyperlipidaemia (familial hypercholesterolaemia)*. London: National Library for Health; 2006. URL: <http://www.library.nhs.uk/screening/ViewResource.aspx?resID=60325&tabID=288>

## Case study 11: Managing the dyspeptic patient: a systematic review and modelling exercise (HTA study 96/37/01)

### Stage 0: Needs assessment

Managing dyspepsia costs the NHS in excess of £500 million per year; 2% of the population consult their GP with dyspepsia each year, and 450,000 endoscopies are performed at a cost of £90 million. Most patients undergoing endoscopy have no significant abnormality and are termed as having non-ulcer dyspepsia (NUD). The vignette for HTA Programme call 96/37 asked for a review of clinical effectiveness and cost-effectiveness of managing the dyspeptic patient.

### Interface A: Project specification and selection

The Principal Investigator, Brendan Delaney, was involved in a trial in this field and through links with other researchers at Birmingham and Leeds Universities put together a team to respond to the call. Their proposal possibly went wider than had been set out in the HTA Programme call and included the initial management of uninvestigated dyspepsia in primary care together with treatments for proven NUD. The study aimed to link systematic reviews with simulation modelling to provide the best available evidence for managing patients with dyspepsia. The main questions asked in relation to the management of uninvestigated dyspepsia in primary care included assessing the effectiveness of: initial pharmacological therapy, early endoscopy, *Helicobacter pylori* screening before endoscopy in patients with dyspepsia and *H. pylori* screening before eradication therapy in patients with dyspepsia. They also attempted to identify both the most cost-effective combinations of initial investigation strategy and prescribing for patients and the most important strategies to compare in future trials. In relation to therapy for proven NUD, they considered the effectiveness of pharmacological therapy and of *H. pylori* eradication, which was the most cost-effective therapy and what were the most important therapies to compare in the treatment of NUD.

### Stage 1: Inputs

The team received £69,019 from the HTA Programme, which paid for a research fellow for each arm of the review and a secretary. There was also some input in kind from other staff members. The HTA funding was seen as particularly important because at that time it was thought unlikely that there would have been alternative funding for such a comprehensive review to be undertaken. In addition to clinical academics with expertise in general practice and in gastroenterology, there was a wide range of other expertise including in health economics and medical statistics.

### Stage 2: Processes

The review was conducted from June 1998 to July 1999 and broadly followed a standard Cochrane Review approach, but with the addition of modelling of the data coming from the review. Such modelling was then fairly rare, but is now an approach more commonly adopted. The Cochrane Collaboration Controlled Trials Register and Database of Systematic Reviews, MEDLINE, EMBASE, CINAHL, SIGLE and ISCI were

searched up to January 1999. Experts in the field of dyspepsia, major pharmaceutical companies and journal editors were also contacted. Authors of publications only available as abstracts were contacted for full trial results. Dyspepsia was defined following the 1988 Rome Working Party definition as any symptom referable to the upper gastrointestinal tract lasting for more than 4 weeks. Two reviewers independently selected eligible trials, according to agreed criteria for dyspepsia in primary care and pharmacological treatment for NUD.

The HTA monograph provides details of how the data were synthesised and of the health economics modelling, which adopted an NHS perspective.

### Stage 3: Primary outputs

#### Knowledge production

In their conclusion, the authors stated: “There is still much uncertainty around the management of dyspepsia, both uninvestigated dyspepsia and proven NUD. This review indicates that the treatment for NUD, for which the evidence is most reliable, is *H. pylori* eradication. The effect is small but cost-effective as the treatment is potentially curative rather than just suppressive. Whether the effect is due to treating latent peptic ulcer disease or some other mechanism, the implication is that patients diagnosed on the basis of a negative endoscopy will benefit from *H. pylori* eradication.

In primary care, the conclusions are much less robust. Proton pump inhibitors (PPIs) are the most effective treatment for undiagnosed dyspepsia and reasonably cost-effective. This is because the case-mix includes patients with peptic ulcer disease and gastro-oesophageal reflux disease, for which PPIs are effective treatments. The relative efficacy of H<sub>2</sub>-receptor antagonists is uncertain, because of a lack of trials comparing antacids and H<sub>2</sub>-receptor antagonists and a lack of trials in patients without reflux as a dominant symptom. Although management based on early endoscopy may lead to a small reduction in dyspeptic symptoms, the cost-effectiveness of endoscopy is uncertain. Modelling suggests that, for most patients, endoscopy-based management is not cost-effective as there is little gain in symptom relief and considerable additional cost. Of the empirical strategies, *H. pylori* test-and-treat is likely to be more cost-effective than endoscopy but well-designed, primary care based trials are needed to compare cost-effectiveness and effects on quality of life with empirical acid suppression.” (Delaney and colleagues, 2000, p. v).

The authors also made various recommendations for both primary research, in the treatment of NUD and for the initial management of dyspepsia, and for secondary research including keeping the reviews up-to-date.

In addition to the rapidly published HTA monograph, a large number of publications arose from this study and also in some cases from the stream of work linked to it (see research targeting below). There were a dozen abstracts, papers in professional journals and contributions to edited works, including *Clinical Evidence*, books and reports. The list below is limited to the monograph, the main articles, and the initial Cochrane Reviews produced by the team who formed a Cochrane Review Group. These were updated at various times as part of the ongoing work and because of problems in identifying specific Cochrane Reviews on the SCI the precise citation figures are not given, but there is evidence they were all being cited. The Cochrane Reviews became more important as they updated and incorporated a growing body of trials and yet still drew substantially on the initial analysis undertaken for the HTA review which is acknowledged in each case as the original source of funding. The *BMJ* article, which was accompanied by an editorial (Veldhuyzen van Zanten, 2000), received the highest number of cites of any publication from any of the case studies and, at minimum, has clearly been taken notice of by a wide range of researchers internationally.

Delaney B, Moayyedi P, Deeks J, Innes M, Soo S, Barton P, *et al.* The management of dyspepsia: a systematic review. *Health Technol Assess* 2000;4(39). [at initial project scoring in 2005, 367 despatched, 13,298 hits, 17 cites; current, 368 despatched, 21,692 hits, 20 cites].

Delaney BC, Innes MA, Deeks J, Wilson S, Cooner MK, Moayyedi P, *et al.* Initial Management strategies for dyspepsia. In *The Cochrane Library*, Issue 3. Oxford: Update Software; 2001.

Delaney BC, O’Morain C. The management of dyspepsia. The year in *Helicobacter pylori*. *Curr Opin Gastroenterol* 2001;17:S38–42 [editorial]. [2004 journal impact factor, 0.5; at scoring, 0 cites, current, 0 cites].

Delaney BC. Dyspepsia management in the millennium: to test and treat or not? *Gut* 2003;52:10–11 [editorial]. [2004 journal impact factor, 6.6; at scoring, 1 cite; current, 2 cites].

Moayyedi P, Soo S, Deeks J, Delaney B, Harris A, Innes M, *et al.* Eradication of *Helicobacter pylori* for non-ulcer dyspepsia. In *The Cochrane Library*, Issue 4. Oxford: Update Software; 2000.

Moayyedi P, Soo S, Deeks J, Foreman D, Mason J, Innes M, *et al.*, on behalf of the Dyspepsia Review Group. Systematic review and economic evaluation of *Helicobacter pylori* eradication treatment for non-ulcer dyspepsia. *BMJ* 2000;**321**:659–64. [2004 journal impact factor, 7; at scoring, 139 cites; current, 148].

Moayyedi P, Soo S, Deeks J, Forman D, Harris A, Innes M, *et al.* Systematic review: antacids, H<sub>2</sub> receptor antagonists, prokinetics, bismuth and sulcralfate therapy for non ulcer dyspepsia. *Aliment Pharm Ther* 2003;**17**:1215–27. [2004 journal impact factor, 3.9; at scoring, 6 cites; current, 12 cites].

Soo S, Moayyedi P, Deeks J, Delaney B, Harris A, Innes M, *et al.* Pharmacological interventions for non-ulcer dyspepsia. In *The Cochrane Library*, Issue 3, Oxford: Update Software; 2001.

Soo S, Moayyedi P, Deeks J, Delaney B, Lewis M, Forman D. Psychological interventions for non-ulcer dyspepsia. In *The Cochrane Library*, Issue 3. Oxford: Update Software; 2001.

### **Research targeting, capacity building and absorption**

The review helped to target further work in several major and overlapping ways. Team members were successful in obtaining £800,000 from the MRC for an RCT to address gaps in the evidence identified by the review; they also proposed using some of the modelling developed for the review. Delaney was also successful in a bid to continue work on this field through obtaining a career scientist award in the 1999 round of the Research Capacity Development's Primary Care Award Scheme. The successful project not only helped advance the careers of some of the key team members but also allowed them to build a Cochrane Review Group that has continued, as they recommended, to build on the original HTA review through undertaking a series of updates to the original review. Furthermore, the original review of *H. pylori* published in the *BMJ* was published about the same time as another review in the USA that had slightly different findings and this led to considerable debate; eventually the two teams worked together on a further review.

### **Interface B: Dissemination**

In addition to many poster displays, the Principal Investigator gave about a dozen oral presentations, including at international academic/practitioner conferences with particularly important ones including an invited keynote address at Digestive Disease Week 2002 in San Francisco. An important and direct way of disseminating the findings was through Delaney's appointment as the technical lead on the team producing guidelines on dyspepsia for NICE (NICE, 2004) (see below). Furthermore, Delaney was able to ensure that some

of the findings were directly disseminated to many practitioners in the UK, and the public, through being invited to be the author of the section on *H. pylori* infection in *Clinical Evidence* (Delaney and colleagues, 2002) and of a similar section in *Best Treatments*, that is published by BMJ Publishing in association with *Clinical Evidence*.

### **Stage 4: Secondary outputs: informing policy**

Not only, as noted, was Delaney the technical lead for the group developing the NICE guidelines, *Dyspepsia: managing dyspepsia in adults in primary care*, but also, as reported in the version of the guidelines that describes the methods used in their development, "The evidence base was derived from published reports, whose review methods are reported comprehensively" (North of England Dyspepsia Guideline Development Group, 2004, p. 34). The two references given for this are the HTA report and the 2000 Cochrane Review by Moayyedi and colleagues. The details of the guidelines therefore draw substantially on the HTA project and the additional work. Similarly, the National Clinical Guideline 68, *Dyspepsia*, from SIGN cites various publications from this stream of work. For example, the *BMJ* article and the Cochrane Review by Moayyedi and colleagues on the eradication of *H. pylori* for non-ulcer dyspepsia are cited as references to support the recommendation that, "*H. pylori* eradication should be considered in the management of functional dyspepsia" (SIGN, 2003, para. 5.4.2). This recommendation is given the highest grade for the strength of evidence on which it is based.

Publications from this stream of work are also well represented in the recent official recommendations of the American Gastroenterologist Association on dyspepsia (AGA, 2005a) and in the accompanying technical review (AGA, 2005b).

### **Stage 5: Applications by practitioners and public**

In contrast to the strong evidence collected in the above sections, this impact assessment identified very little evidence about the application of the findings by practitioners and the public. Team members certainly produced papers/summaries in publications known to be targeted at practitioners and the public in both the UK and the USA.

### **Stage 6: Impacts or final outcomes**

Again, little evidence was identified about the final outcomes, although the evidence has become even stronger about *H. pylori* eradication being cost-effective whereas endoscopy was not.

## Comments

The ability of the research team to widen the scope of the review somewhat at the project specification stage contributed to the considerable impact made by the project. In conducting the assessment, the well-completed questionnaire and the detailed interview were most useful in tracing the impact. Furthermore, they also provided sources for further documents that it was helpful to access and analyse. The project provides an almost classic example of a systematic review that was cost-effective research in that it became the basis of a series of Cochrane Reviews and helped target further well-funded trials and it informed key policy documents. The only negative element of the case study was the inability to make much progress in measuring impact on practice and subsequent outcomes.

## References

- American Gastroenterologist Association. American Gastroenterologist Association medical position statement: evaluation of dyspepsia. *Gastroenterology* 2005a;**129**:1753–5.
- American Gastroenterologist Association. American Gastroenterologist Association technical review on the evaluation of dyspepsia. *Gastroenterology* 2005b;**129**:1756–80.
- Delaney B, Moayyedi P, Deeks J, Innes M, Soo S, Barton P, *et al.* The management of dyspepsia: a systematic review. *Health Technol Assess* 2000;**4**(39).
- Delaney B, Moayyedi P, Forman D. *Helicobacter pylori* infection. *Clin Evid* 2002;**9**:0–2.
- National Institute for Clinical Excellence. *Dyspepsia: managing dyspepsia in adults in primary care*. Evidence-based Clinical Practice Guideline. London: NICE; 2004.
- North of England Dyspepsia Guideline Development Group. *Dyspepsia: managing dyspepsia in adults in primary care*. Centre for Health Services Research Report No. 112. Newcastle: CHSR, University of Newcastle upon Tyne; 2004.
- Scottish Intercollegiate Guidelines Network. *Dyspepsia*. National Clinical Guideline No. 68. Edinburgh: Royal College of Physicians; 2003.
- Veldhuyzen van Zanten SJO. Treating non-ulcer dyspepsia and *H pylori*. (Editorial). *BMJ* 2000;**321**:648–9.

## Case study 12: Systematic review and evaluation of methods of assessing urinary incontinence (HTA study 99/29/02)

### Stage 0: Needs assessment

Urinary incontinence is not life threatening, but it can have enormous costs to individuals and the

health service in terms of expenditure and impact on quality of life. Epidemiological studies have demonstrated that urinary incontinence is a very common symptom, with a reported prevalence of any urinary incontinence (in those aged 40 years and over) of 34% for women and 14% for men.

Pathways to diagnostic assessment are inconsistent, with some individuals being assessed and treated in primary care settings by GPs and nurses, and others being referred directly to a variety of specialists in secondary care (e.g. physiotherapists, gynaecologists and urologists) without any assessment or treatment. Assessment can be undertaken at a number of levels using different combinations of tests.

It is particularly important when implementing certain treatment interventions (e.g. medication that may have side-effects) that a diagnosis is made to determine the most effective treatment intervention, and it is imperative before surgical intervention. If a diagnosis is not made, then inappropriate and unnecessary interventions may be implemented. Two types of diagnosis can be made: symptomatic diagnosis and condition-specific diagnosis. The evidence available on the accuracy and acceptability of these diagnostic processes is inconsistent and variable.

The HTA Programme call 99/29 set out that there was growing evidence that appropriate management could reduce the morbidity and cost of urinary incontinence, but there was variable use of clinical and diagnostic assessment. Therefore, there was a need to assess the cost-effectiveness of different methods of diagnostic assessment.

### Interface A: Project specification and selection

The Principal Investigator, Kate Williams of the Department of Health Sciences at the University of Leicester, had been involved in various trials related to incontinence. In response to the HTA call, she brought together a team including researchers experienced in conducting systematic reviews and in incontinence research. The proposal set out that the search strategy would be based on the Cochrane and NHS Centre for Reviews and Dissemination strategies for identifying studies of diagnostic performance. No major revisions were requested from peer reviewers. The objectives of the review as set out in the HTA monograph were to:

- identify, appraise and summarise the published evidence relating to different methods of



diagnostic assessment of male and female urinary incontinence, specifically urodynamic stress incontinence (USI) and detrusor overactivity (DO)

- quantitatively synthesise the extracted evidence using meta-analysis methods (where possible) or pooling of individual sensitivity and specificity data
- construct an economic model to examine the cost-effectiveness of simple, commonly used primary care tests
- identify gaps in the literature and prioritise future clinical and research questions. (Martin and colleagues, 2006, p. ix).

### Stage 1: Inputs

The project received £69,893 from the HTA Programme and received no other funding. The funding covered the employment of the research fellow. The research team included expertise from a wide range of disciplines, including urology, nursing, gynaecology, statistics and health economics.

### Stage 2: Processes

There were some delays due to the maternity leave of the Principal Investigator and the study ran from April 2002 to June 2003. Although the planned approach broadly worked, there was comparatively little good research identified, despite the comprehensive searches. In the review, the online bibliographic databases MEDLINE (1966–2002), CINAHL (1982–2002) and EMBASE (1980–2002) were used to obtain the literature. Study selection comprised a three-stage process using defined inclusion and exclusion criteria. All records were assessed for relevance by the first investigator on the basis of the abstract, or if the abstract was not available then title only. Papers were considered relevant to the systematic review if they considered the evaluation, appropriateness and/or cost of diagnostic assessment in the following categories: clinical history-taking; simple investigations including validated scales, diaries and pad tests; and advanced (invasive) investigations (e.g. urodynamics). To be included, a paper had to provide a quantitative comparison between two or more different methods of diagnosing urinary incontinence.

A panel consisting of at least three members of the review team, including at least one statistician, discussed all papers identified as of potential relevance. The panel determined whether study data were presented in a suitable format to calculate sensitivity and specificity. All relevant

papers were assessed for quality using Quality Assessment of Diagnostic Studies (QUADAS), a tool designed specifically for studies on diagnostic accuracy. An initial pilot study on four papers resulted in some clarifications being added to the instructions of the QUADAS tool to ensure consistency between assessors. Seven of the authors performed the full quality assessment process, with 10% of the papers being assessed by two authors to test for inter-reader agreement.

Studies that reported the results of applying the same diagnostic procedure using the same threshold value (cut-off) were pooled using a random effects meta-analysis model to produce pooled estimates of sensitivity, specificity and diagnostic odds ratio together with 95% confidence intervals (Martin and colleagues, 2006, pp. ix–x).

### Stage 3: Primary outputs

#### *Knowledge production*

The HTA monograph was submitted in November 2003 but not published until February 2006, which, as discussed in the comments section below, could be a cause for concern. The main results were that, “In total, 6009 papers were identified from the literature search, of which 129 were deemed relevant for inclusion in the review, and these papers compared two or more diagnostic techniques. The gold-standard diagnostic test for urinary incontinence with which each reference test was compared was multichannel urodynamics.

“In general, reporting in the primary studies was poor; there was a lack of literature in the key clinical areas and minimal literature dealing with diagnosis in men. Only a limited number of studies could be combined or synthesised, providing the following results when compared with multichannel urodynamics. A clinical history for diagnosing USI in women was found to have a sensitivity of 0.92 and specificity of 0.56 and for DO a sensitivity of 0.61 and specificity of 0.87. For validated scales, question 3 of the Urogenital Distress Inventory was found to have a sensitivity of 0.88 and specificity of 0.60. Seven studies compared a pad test with multichannel urodynamics; however, four different pad tests were studied and therefore it was difficult to draw any conclusions about diagnostic accuracy.

“Of the four studies comparing urinary diary with multichannel urodynamics, only one presented data in a format that allowed sensitivity and specificity to be calculated. Their reported values of 0.88 and 0.83 suggest that a urinary diary may

be effective in the diagnosis of DO in women. Examination of the incremental cost-effectiveness of three primary care tests used in addition to history found that the diary had the lowest cost-effectiveness ratio of between £35 and £77 per extra unit of effectiveness (or case diagnosed). Imaging by ultrasound to determine leakage was found to be effective in the diagnosis of USI in women, with a sensitivity of 0.94 and specificity of 0.83.” (Martin and colleagues, 2006, p. x).

The authors noted that this was the first systematic review of methods for diagnosing urinary incontinence, but that, “As reporting of the primary studies was poor, clinical interpretation was often difficult because few studies could be synthesised and conclusions made.” (Martin and colleagues, 2006, p. x). They nevertheless suggested that: a large proportion of women with USI can be correctly diagnosed in primary care from clinical history alone; on the basis of diagnosis the diary appears to be the most cost-effective of the three primary care tests used; ultrasound imaging may offer a valuable alternative to urodynamic investigation; the clinical stress test is effective in the diagnosis of USI; and if a patient is to undergo an invasive urodynamic procedure, multichannel urodynamics is likely to give the most accurate result in a secondary care setting.

The team waited until the final acceptance of the HTA report before submitting articles which caused considerable delay, but since publication it has already attracted a fairly large number of hits:

Martin JL, Williams KS, Abrams KR, Turner DA, Sutton AJ, Chapple C, *et al.* Systematic review and evaluation of methods of assessing urinary incontinence. *Health Technol Assess* 2006;**10**(6). [not published at scoring; current, 119 despatched, 8509 hits, no cites].

#### **Research targeting, capacity building and absorption**

The fact that the work has only just been published reduces the possibilities for having had an impact on future research. Nevertheless, the emphasis that the study gives to the limited nature of the existing stock of good-quality studies could possibly encourage the funding of more primary research in this area, especially in relation to males. The study led to no further qualifications being gained but possibly helped the successful career development of the research fellow.

#### **Interface B: Dissemination**

Several presentations were given by the Principal Investigator and by other members of the team to

academic and to practitioner audiences, including internationally, and they generated some interest.

#### **Stage 4: Secondary outputs: informing policy**

Although it is too early to have expected much impact, the study was cited as work in progress in the chapter on diagnostics in a report from the 3rd International Consultation on Incontinence held in 2004 (Staskin and colleagues, 2005). This early impact might be taken to indicate the possibility for further such impact.

#### **Stage 5: Applications by practitioners and public**

A small number of clinicians responded to the presentations by saying to the Principal Investigator that they might adopt some of the recommendations such as using the stress test for assessing stress incontinence. Again, this might indicate the potential for further impact in the future.

#### **Stage 6: Impacts or final outcomes**

If the recommendations from the study were adopted, this could result in an improved quality of health service delivery for some patients, and perhaps some cost savings, as a result of the less invasive and cost-effective diagnostic procedures being more widely adopted for some patients. They may provide sufficient information to commence primary care interventions which are low risk and low cost.

#### **Comments**

In any impact assessment, there is often less to say and to explore in relation to a recently completed study such as this. The delays between the submission of the report by the authors in November 2003 and its publication in February 2006 might, however, be a cause for concern and potentially reduce the study’s impact. The number of hits that the report has attracted in the 7 months since publication indicates that it was a study/topic in which there was considerable interest. The researchers suggested that the NCCHTA could consider identifying potential reviewers much earlier in the process to determine whether they would have the necessary time available to review it once it had been submitted (this is something that the NCCHTA has started routinely doing in the last 2 years).

#### **References**

Martin JL, Williams KS, Abrams KR, Turner DA, Sutton AJ, Chapple C, *et al.* Systematic review and evaluation of methods of assessing urinary incontinence. *Health Technol Assess* 2006;**10**(6).

Staskin D, Hilton P, Emmanuel A, Goode P, Mills I, Shull B, *et al.* Initial assessment of incontinence. In Abrams P, Cardozo L, Khoury S, Wein A, editors. *Incontinence*. Plymouth: Health Publications; 2005.

### **Case study 13: A systematic review to examine the impact of psycho-educational interventions on health outcomes and costs in adults and children with asthma (HTA 01/16/02)**

#### **Stage 0: Needs assessment**

As set out in the HTA Programme call 01/16 and the HTA monograph (Smith and colleagues, 2005), this project focused on a particularly complex issue. Despite effective treatments and management guidelines, there are a significant minority of asthma patients who suffer from severe or poorly controlled disease. When persistent, this is sometimes referred to as 'difficult' asthma, or in some cases 'brittle asthma'. Research highlights the association of psychosocial factors with difficult asthma and its related adverse consequences (e.g. fatal and near-fatal attacks). Existing reviews of programmes involving interactive education, training in self-management and/or targeting specific psycho-social issues resulting from or impacting on asthma suggested that some psycho-educational interventions were effective and potentially cost-effective in general asthma populations. However, it was thought that such findings were unlikely to be generalisable to patients with difficult asthma, hence the HTA Programme identified that a systematic review was required to evaluate the benefits and cost-effectiveness of psychological interventions for managing asthma not controlled by standard treatment.

#### **Interface A: Project specification and selection**

A team of researchers at East Anglia University were conducting a trial related to the psychological aspects of asthma. When the HTA Programme call 01/16 was launched, it fitted closely with issues being addressed in the trial. The principal investigator, Miranda Mugford, encouraged and very strongly supported by the lead researcher on the trial, Jane Smith, constructed a team from the local medical school and hospital who had experience in research into asthma, psychology, systematic reviewing and health economics. As a result of the review process, a local GP and a paediatric specialist were added to the team. A

thorough review was proposed to provide a broad overview of all potentially relevant literature and to address the questions of whether psycho-educational interventions improve outcomes for patients with difficult asthma and whether they constitute an efficient use of healthcare resources.

#### **Stage 1: Inputs**

The project received £80,282 from the HTA Programme and there was no other funding apart from the time provided by the senior faculty. The funding covered 10% of the time of the main researcher, a further full-time researcher and 50% of the time of an administrator. Staff from several faculties contributed time for reading and reviewing papers. The Principal Investigator had considerable experience of HTA projects, including providing the economic evaluation in a number of trials (see, for example, case study 6 above).

#### **Stage 2: Processes**

The review was conducted from February 2002 to January 2003. It was in an exploratory area and therefore, rather than starting with a protocol with clear criteria, it was first necessary to conduct an exercise to develop appropriate definitions. As described in the monograph, the researchers combined asthma terms with complex permutations for describing interventions. These were used to search 32 electronic data sources (including research registers, grey literature and non-English language databases) and guide handsearching of reference lists, conference proceedings, *Current Contents* and three key journals up to the end of 2002.

Abstracts and/or titles were assessed in duplicate, against the definitions developed at the start of the review, to identify potentially eligible interventions targeting patients with forms of or one or more risk factors/outcomes associated with difficult asthma. Final inclusion decisions were made on the basis of viewing full texts. Two reviewers classified the studies initially included by patient group (child, adult) and graded them along dimensions related to study design and relevance in terms of the degree to which they were judged to target difficult asthma (insufficient, possible, probable, definite). A third reviewer resolved disagreements or uncertainties.

Descriptive, methodological, outcome and cost data were extracted from studies meeting a minimum design (having a control group) and relevance (at least 'possible' targeting of difficult

asthma) threshold. Authors were contacted for additional information as necessary.

Characteristics of studies in children and adults selected for in-depth review were tabulated separately and results qualitatively synthesised. Where sufficiently similar studies reported adequate data about comparable outcomes, quantitative syntheses (meta-analyses) of results were undertaken using a random effects approach to calculate pooled relative risks (RRs), or standardised mean differences (SMDs) with 95% confidence intervals (CIs) (Smith and colleagues, 2005).

### Stage 3: Primary outputs

#### Knowledge production

There was a considerable time-lag between the completion of the review and the publication of the monograph in June 2005. The main findings set out in the monograph are that, "From over 23,000 citations identified, 4240 abstracts and/or titles were considered for further review. A total of 278 citations reporting on 188 different studies were initially included and classified. Of these, 57 (35 in children, 21 in adults and one including child and adult subgroups) were considered suitable for in-depth review. ...

"The delivery, setting, timing and content of interventions varied considerably even within broad types. Reporting of interventions and methodological quality was often poor but studies demonstrated some success in targeting and following up at-risk patients. The range of outcomes assessed and variations in the ways they were measured and reported precluded quantitative synthesis for most. ...

"There was evidence that, compared with usual or non-psycho-educational care, psycho-educational interventions reduced admissions when data from the latest follow-ups reported were pooled across nine studies in children (RR = 0.64, 95% CI = 0.46 to 0.89) and six studies with possible targeting of difficult asthma in adults (RR = 0.57, 95% CI = 0.34 to 0.93). In children, the greatest and only significant effects were confined to individual studies with limited targeting of difficult asthma and no long-term follow-up. Limited data in adults also suggested that effects may not extend to those most at risk. There was no evidence of pooled effects of psycho-educational interventions on emergency attendances from eight studies in children (RR = 0.97, 95% CI = 0.78 to 1.21) and four in adults (RR = 1.03, 95% CI = 0.82 to 1.29).

"There were overall significant reductions in symptoms, similar in different subgroups of

difficult asthma, across four paediatric studies that could be combined (SMD = -0.45, 95% CI = -0.68 to -0.22), but mixed results across individual adult studies. A small number of individual studies in children showed mainly positive effects on measures of self-care behaviour but, with respect to all other outcomes in adults and children where sufficient data allowed conclusions to be drawn, studies showed mixed results or suggested limited effectiveness of psycho-educational interventions. No studies of psychosocial interventions were included in any quantitative syntheses and it was not possible to draw clear conclusions regarding the relative effectiveness of educational, self-management and multifaceted programmes. ...

"Data on costs were very limited in quantity and quality for children and adults. Of the two well-designed economic evaluations identified, both of multifaceted interventions, one in children suggested that, from the health provider's viewpoint, there would be an additional cost of achieving health gain in terms of symptom-free days. Provisional data from the other study suggested that in adults the significantly increased costs of providing an intervention were not offset by any short-term savings in use of healthcare resources or associated with improvements in health outcomes." (Smith and colleagues, 2005, pp. ix-x).

Several relevant well-designed UK studies assessing cost-effectiveness were not published in time to be included, but the main researcher has continued to gather and analyse papers so that they could be referred to in articles from the study. For example, a paper was submitted in summer 2006 that updated the adult part of the review.

The authors concluded that although there was some evidence of overall positive effects of psycho-educational interventions on hospital admissions in adults and children, and on symptoms in children, there was limited evidence of effects on other outcomes. They therefore stated: "There is currently a lack of evidence to warrant significant changes in clinical practice with regard to care of patients with more difficult asthma." (Smith and colleagues, 2006, p. xi). They listed a range of key areas for further research, both primary and secondary.

Various presentations and poster abstracts from national and international conferences have been published in addition to the HTA report, which has received a considerable number of hits in the year since it was published:

Smith JR, Mugford M, Holland R, Candy B, Noble MJ, Harrison BDW, *et al.* A systematic review to examine the impact of psycho-educational interventions on health

outcomes and costs in adults and children with difficult asthma. *Health Technol Assess* 2005;9(23). [at initial project scoring in 2005, 63 despatched, 1482 hits, no cites; current, 116 despatched, 8010 hits, 1 cite].

### **Research targeting, capacity building and absorption**

The HTA report set out a comprehensive list of areas that would benefit from further research and provided some priorities for this field, but it remains a very complex topic; for example, it is not clear how far the main interventions should be medical, psychological or social. There has not been much time for the research to impact on the agenda of other researchers, but members of the HTA team conducted some further work informed by the review, and the earlier trial, and this led to applications for further funding. Smith published a paper with Mike Noble on use of GP registers to identify at-risk patients (Noble and colleagues, 2006). In addition, funding has been granted by Asthma UK to a team at UEA led by Smith and Andrew Wilson, and including Mugford on the economic analysis, to work on a larger study to evaluate care based on use of primary care asthma risk registers to prompt caregivers. Both of these fill some of the gaps identified in the review (i.e. that some 'difficult asthma' was actually due to failures in provision or access to medical care, despite the definition of the condition being 'asthma unresponsive in spite of best medical care', and that there was a need for opportunistic interventions based in primary care). Furthermore, mainly through increasing Mugford's knowledge of the literature on the economics of allergy, the HTA study also indirectly fed into a large project on food allergies, funded under the Sixth European Framework, for which she is leading on the economic analysis.

The main researcher has been promoted to Lecturer in Health Psychology and is completing a PhD that draws on this study, and the study has also made some contribution to the career development of the other members of research staff. The review has possibly helped to bring some of the international research to the attention of UK audiences and also, through the quality assessments, helped to identify limitations in some of the studies making strong, but not well founded, claims of cost-effectiveness.

### **Interface B: Dissemination**

As noted above, there has been a considerable effort to disseminate both the findings and the methodological issues raised when undertaking a review in a field such as this. There have been at least five presentations to national and international academic audiences and three to practitioners.

### **Stage 4: Secondary outputs: informing policy**

The scope for impact on policy from this review is limited, but it could potentially feed into guidelines on this topic.

### **Stage 5: Applications by practitioners and public**

Again, there is limited scope for impact, but the review could have had some impact on practitioners at the local level given the interest they have shown in the area.

### **Stage 6: Impacts or final outcomes**

There would be a potential for some health gain through improved symptoms following the provision of psycho-educational interventions, but this would be at a potentially increased overall cost and thus far the interventions seem unlikely to produce the cost-effectiveness claimed by some studies in the USA.

### **Comments**

This case study illustrates that it can be necessary to allow considerable time to elapse after the completion of a project before a reasonable assessment of impact can be undertaken, but that the picture in terms of hits on the URL of the HTA monograph can change quite dramatically over a short period. This rise in the number of hits indicates a considerable level of interest in the findings. The study also illustrates how the key role of some reviews is likely to be the systematic identification of gaps in the evidence rather than the provision of clear answers to the initial questions.

### **References**

- Noble MJ, Smith JR, Windley J. A controlled retrospective pilot study of an 'at-risk asthma register' in primary care. *Prim Care Respir J* 2006;15:116–24.
- Smith JR, Mugford M, Holland R, Candy B, Noble MJ, Harrison BDW, *et al.* A systematic review to examine the impact of psycho-educational interventions on health outcomes and costs in adults and children with difficult asthma. *Health Technol Assess* 2005;9(23).

## **Case study 14: The clinical effectiveness and cost-effectiveness of riluzole for the treatment of motor neurone disease (NICE TAR 00/01/01)**

### **Stage 0: Needs assessment**

For this TAR, NICE identified a need for an assessment of the clinical effectiveness and cost-effectiveness of riluzole (trade name Rilutek<sup>®</sup>), a

drug used to treat people with the amyotrophic lateral sclerosis (ALS) form of motor neurone disease (MND). Its licensed indication is to extend life or the time to mechanical ventilation, and it costs about £3700 per year. The prevalence of MND is approximately seven per 100,000, and ALS constitutes 65–85% of this. Incidence rises with age. The disease is extremely distressing for patients and their carers, and is relentlessly progressive, with death usually occurring within 3–5 years. Death usually occurs from respiratory infection and failure, and complications of immobility. There is no cure and treatment consists mostly of symptomatic, supportive and palliative care.

### Interface A: Project specification and selection

NICE set the scope or objective of the assessment as being a relatively straightforward clinical effectiveness and cost-effectiveness study. This TAR was commissioned from the West Midlands Development and Evaluation Service based at the University of Birmingham.

#### Stage 1: Inputs

The NICE TAR Centres work under a contract which involves them conducting a number of TARs each year on a ‘call-off contract basis’ without a precise specification of the amount of funding for each one. The average ‘cost’ of TARs is discussed elsewhere in this report, but at this time a TAR was considered to cost about £40,000. All TAR teams consist of researchers with experience of conducting health technology assessments, and one of the team worked almost full-time on it with others making a contribution. As usual, helpful advice was provided through consultation with an advisory group of experts.

#### Stage 2: Processes

The study was originally conducted in 3 months between May and August 2000, but the actual work took a lot longer because of the new data received after the report was first submitted. As reported in the HTA monograph, a systematic review of RCTs and economic studies addressing the clinical effectiveness and cost-effectiveness of riluzole in MND was undertaken. Electronic databases, reference lists from publications, conference abstracts and the Aventis Pharma submission to NICE were searched. Clinical experts and specialist organisations were also contacted. Studies were included if they had investigated either clinical effectiveness, cost-effectiveness or safety of riluzole, or quality of

life/patient satisfaction associated with its use in MND patients, with no restrictions on age or sex.

The review adhered to the guidance of the West Midlands Development and Evaluation Service Handbook and the York Centre for Reviews and Dissemination guidelines. The economic analysis involved a rigorous assessment of the strengths and weaknesses of existing analyses, and built a further model to explore the impact uncertainties revealed. Although it had not been in the original specification, the authors thought that it was important to include in the review a chapter, entitled ‘Patient perspectives’, drawing on more qualitative data in published documents, including a Danish study entitled *Between hope and despair* (Danish Institute for Health Services Research and Development, 1998). Finally, after the TAR had been submitted, further data were provided by the manufacturers of riluzole, and an update was supplied to NICE and subsequently incorporated into the HTA monograph.

#### Stage 3: Primary outputs

##### Knowledge production

The main findings were reported in the TAR and subsequent HTA monograph. Four studies met the inclusion criteria for the clinical effectiveness review. All compared riluzole with placebo and reported tracheostomy-free survival as a main outcome.

In the Executive Summary of the HTA monograph, which incorporated findings from the update, the authors reported, “Median follow-up in all trials was 18 months with most patients having follow-up of between 16 and 21 months. Combined results favoured riluzole with a hazard ratio for tracheostomy-free survival of 0.88 (95% confidence interval (CI), 0.75 to 1.02). ... Riluzole does not improve symptoms. When data on functional status were combined, a small reduction in the rate of deterioration of functional status was observed, although it was not clear whether this was clinically significant. A large proportion of patients in both groups reported adverse events, but there was little overall difference between riluzole and placebo.

“... It is clear that riluzole is associated with a net increase in costs to the health service, although the magnitude of the increase is difficult to predict accurately. A more robust estimate of the riluzole-induced gain in life expectancy over the whole disease duration is required to diminish current uncertainties relating to methods of extrapolation beyond observed survival in trials. In our model,

base-case incremental cost-effectiveness ratio (ICER) produced a cost per life-year of £39,000 and a cost per quality-adjusted life-year (QALY) of £58,000. A sensitivity analysis indicated that the most optimistic ICER (cost per QALY) is £20,000 and the most pessimistic has riluzole dominated by placebo.” (Stewart and colleagues, 2001, p. iv).

The authors concluded, “There is limited evidence of a modest benefit in tracheostomy-free survival for patients taking riluzole. However, the evidence is restricted and uncertainty remains as to the true benefit of riluzole; the CI is wide and compatible with little or no difference between riluzole and placebo. When costs and the health economic impact are considered when extrapolating survival beyond that observed in trials, the uncertainty about whether the benefits are worth the costs is magnified. Even under the most optimistic assumptions, riluzole at best only postpones death for a few months, and does not preclude the need for supportive care and practical help.

“If riluzole were to be made available to all patients in whom it is not contraindicated, the annual cost to the NHS would be about £8.4 million, assuming all these patients wish to take it. Many patients, given accurate information about the benefits and effects of riluzole, may choose not to.” (Stewart and colleagues, 2001, p. iv).

One element of the update was based on newly supplied information that had been used in a Markov model for an earlier published economic analysis (Tavakoli and colleagues, 1999) and also in the economic analysis reported in the company’s submission to NICE (Aventis Pharma). In the update, the authors stated, “We have identified a number of errors (mostly minor, but one serious) in the description and analysis of the model presented to us.” (Stewart and colleagues, 2001, p. 83).

The HTA monograph was the main publication, but there was also a letter in the *BMJ* that is discussed below and a discussion of some of the methodological issues in a book chapter.

Bryan S, Sandercock J, Barton P, Burls A. Tensions in licensing and reimbursement decisions: the case of riluzole for amyotrophic lateral sclerosis. In Freemantle N, Hill S, editors. *Evaluating Pharmaceuticals for Health Policy and Reimbursement*. London: BMJ Books: 2004.

Sandercock J, Burls A, Hyde C, Fry-Smith A, Barton PM, Bryan S, *et al.* Riluzole for motor neurone disease – more trials are needed. *BMJ* 2001;**322**:1305. [letter: at scoring, 5 cites; now 5 cites].

Stewart A, Sandercock J, Bryan S, Hyde C, Barton PM, Fry-Smith A, *et al.* The clinical effectiveness and cost-effectiveness of riluzole for motor neurone disease: a rapid and systematic review. *Health Technol Assess* 2001; **5**(2). [at time of scoring in 2005, 214 despatched, 8608 hits, 6 cites; current, 214 despatched, 13,355 hits, 8 cites].

### **Research targeting, capacity building and absorption**

Although the team has not conducted any further research on riluzole, the methodological analysis has informed further work and teaching of the team both in health economics and in relation to patient preferences. Based partly on this work, team members were funded by NICE for about £31,000 to conduct the Patient Impact Assessment Project to develop ideas about how patient preferences could be incorporated into TARs. Furthermore, it is suggested that the data were also used by others to calculate the expected value of perfect information in this area as part of the methodological work on this topic. The analysis that identified the flaw in the Markov model cited and used by the manufacturers could also feed into further research.

The work is also making some contribution to a team member’s PhD. Illustrating how conducting research in a country can play a role in helping that system to absorb the findings of research conducted elsewhere, the review increased the amount of research in this field that was in the public domain in the UK.

### **Interface B: Dissemination**

The whole structure of the NICE appraisal process gives the researchers a direct feed into the NICE appraisal policy-making process. The TAR was submitted to the Appraisal Committee and team members were invited to a meeting of the committee. In addition, several presentations were made.

### **Stage 4: Secondary outputs: informing policy**

The Technology Appraisal Guidance No. 20 recommends the use of riluzole and the TAR, including the updated version, is listed first in the evidence made available to the Appraisal Committee (NICE, 2001). The press release accompanying the launch of the Guidance referred to it as a clinically and cost-effective intervention. A debate occurred some months later in the *BMJ* about the role of the TAR and how far the guidance was in line with the evidence from the TAR. The authors of the TAR stated, “The evidence in favour of riluzole for use in the

amyotrophic lateral sclerosis form of motor neurone disease is very weak” (Sandercock and colleagues, 2001, p. 1305). Others, including an adviser to the NICE TAR, went further and claimed, “the NICE recommendation is contrary to the conclusion of the expert group at the West Midlands Development and Evaluation Service” (Wheatley and Gray, 2001, p. 1305). In response, the Chairman of the Appraisal Committee stated: “one form of evidence used in an appraisal is a formal systematic review. The assessment report does not make recommendations on how the technology should be used in the NHS; that is the job of the institute’s appraisal committee. The committee also receives submissions from patient and professional organisations, which provide perspectives not captured by a formal review of published evidence” (Barnett, 2001, p. 573).

Unlike the Executive Summaries in the other HTA monographs, those in NICE TARs do not have a section in the conclusions on ‘Implications for healthcare’. Nevertheless, in the discussion section of this TAR, a most interesting argument is made: “The evidence on effectiveness and health economic impact does not unequivocally indicate the use of riluzole in ALS as the best policy for the NHS. However, policy makers may wish to take into account the fact that riluzole is the only specific treatment currently available for ALS” (Stewart and colleagues, 2001, p. 40). The second half of this statement reflects some of the analysis in the *Patient Perspectives* chapter of the report.

Although there is clearly room for differing interpretations, the above analysis does suggest that the NICE TAR had at least some influence on the policy making by NICE and that perhaps the totality of the report and the NICE Guidance are compatible. Indeed, this Guidance was one of a very small number where the Appraisal Committee explicitly used the freedom given to them to consider the values of patients, which mainly come in from the NICE consultation with patient groups, but it is possible to argue that this was strengthened by the points made in that section of the TAR. According to Devlin and Parkin (2004), who analysed the guidance issued by NICE for 39 technologies: “The clearest indication of factors other than cost effectiveness influencing its deliberations is provided by the guidance for Riluzole for MND decision 20)” (p. 443).

In their book chapter discussing the evidence and the decision in relation to riluzole, some of the

original HTA team recognise that in addition to cost-effectiveness considerations, “There are a number of additional factors that may have influenced the reimbursement decision in this case.” (Bryan and colleagues, 2004, p. 134). These factors include both the fact that the overall financial impact of the positive reimbursement decision is likely to be very small, and the equity considerations linked to the lack of alternative treatments for this devastating disease: “there were intangible benefits from being able to do something rather than nothing” (p. 135). Furthermore, they noted that it was very unlikely that there would be any further research addressing the role of riluzole in such a rare disease and so there was little hope that the uncertainties could be reduced. This view would seem to be supported by the standard reviews of the Guidance in 2004 and 2006 which indicated that there was insufficient new evidence to justify a full review appraisal. In April 2006, the Guidance was stated to be ‘static’.

As noted above, the findings in the TAR are used in teaching by team members.

### **Stage 5: Applications by practitioners and public**

Although it is unlikely that the TAR will have had much direct impact on practitioners on its own, to the extent that it did influence the NICE Guidance then it is reasonable to consider how far that has had an impact on the adoption or application of riluzole by practitioners. NICE commissioned Abacus to measure the impact of 28 examples of its Guidance. In some cases the report attempted to address the counter-factual issue by considering the trend line for adopting the particular adoption and compare it with any change in the actual adoption rate following the NICE Guidance. In the case of riluzole, Abacus claim that the trend of increased use of riluzole in secondary care did not change following the Guidance, but that “In primary care, the effect of NICE guidance was to accelerate the use of riluzole, adding about 4,000 units per year” (Abacus International, 2005, p. 9).

### **Stage 6: Impacts or final outcomes**

The health gains from the use of riluzole take the form, as noted above, of the limited evidence for a modest increase in tracheostomy-free survival for patients. In addition, however, there is some evidence of an intangible benefit, perhaps in the form of increased satisfaction with the healthcare provided, if the option of receiving riluzole is made available to patients as it is currently the



only treatment. The NICE Guidance removed 'post-code' prescribing for this particular medication and it could be argued that there is a small equity gain if the only possible treatment is made available.

## Comments

In addition to the publications and the questionnaire, there is considerable written material available on this topic available through review of the NICE website, including the Abacus report, and the citations to the various publications that can be identified through the SCI and *Google Scholar*. Nevertheless, the interview was useful in helping to organise the analysis of the material in this particularly complicated case study.

The circumstances surrounding the policy decision in this case are clearly open to differing interpretations, but several points deserve to be highlighted. First, the role played by this TAR in relation to the analysis of the flaw in the Markov model shows a key strength of the NICE process: it involves independent review of the scientific evidence. Second, the decision in this case is widely viewed as being one in which factors other than cost-effectiveness were taken into account, and in this regard the innovative chapter in the TAR on patient perspectives could have been important in supporting the evidence directly supplied by patient groups. Third, although there remains uncertainty about the precise role of the TAR in informing the decision of the Appraisal Committee, and the extent of the final outcomes is limited, this case study is able to go further than most in providing an estimate of the level of impact on practice – through using the figures from Abacus.

## References

- Abacus International. *NICE Guidance implementation tracking: data sources, methodology, and results*. Abacus International; 2005. URL: <http://www.nice.org.uk/page.aspx?o=209966>
- Barnett D. Riluzole for motor neurone disease. Reply from chairman of appraisal committee at NICE. (Letter). *BMJ* 2001;**323**:573.
- Bryan S, Sandercock J, Barton P, Burls A. Tensions in licensing and reimbursement decisions: the case of riluzole for amyotrophic lateral sclerosis. In Freemantle N, Hill S, editors. *Evaluating Pharmaceuticals for health policy and reimbursement*. London: BMJ Books; 2004.
- Danish Institute for Health Services Research and Development. *Between hope and despair: ALS patients and riluzole*. DSI Rapport 98.03. Copenhagen: Danish Institute for Health Services Research and Development; 1998.

Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Econ* 2004;**13**:437–52.

National Institute for Clinical Excellence. *Guidance on the use of riluzole (Rilutek) for the treatment of motor neurone disease*. Technology Appraisal Guidance No. 20. London: NICE; 2001.

Sandercock J, Burls A, Hyde C, Fry-Smith A, Barton PM, Bryan S, *et al*. Riluzole for motor neurone disease. (Letter). *BMJ* 2001;**322**:1305.

Stewart A, Sandercock J, Bryan S, Hyde C, Barton PM, Fry-Smith A, *et al*. The clinical effectiveness and cost-effectiveness of riluzole for motor neurone disease: a rapid and systematic review. *Health Technol Assess* 2001;**5**(2).

Tavakoli M, Davies HTO, Malek M. Modelling the long-term cost-effectiveness of riluzole for the treatment of amyotrophic lateral sclerosis. *J Drug Assess* 1999; **2**:219–32.

Wheatley K, Gray R. Large, more economical trial is needed. (Letter). *BMJ* 2001;**322**:1305.

## Case study 15: 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 (NICE TAR 00/13/01)

### Stage 0: Needs assessment

For this TAR, NICE identified the need to review the clinical effectiveness and cost-effectiveness evidence about three possible therapies for advanced colorectal cancer: irinotecan, oxaliplatin and raltitrexed. Colorectal cancer is the second most common cause of cancer death in the UK, causing almost 15,000 deaths in England and Wales in 1998. In 1992, a total of 29,664 new cases were registered in England and Wales, an incidence of 56.6 per 100,000 population. Advanced colorectal cancer has been defined as colorectal cancer that, at presentation or recurrence, is either metastatic or so locally advanced that surgical resection is unlikely to be carried out with curative intent. About 80% of patients diagnosed with colorectal cancer undergo surgery. Many have potentially good survival outcomes following surgery (with adjuvant chemotherapy in some cases), but over 50% of patients who have undergone surgery with apparently complete excision will eventually develop advanced disease and distant metastasis (typically presenting within 2 years of initial

diagnosis). Median survival from diagnosis of metastatic disease is 6–9 months, and during this time patients may develop a wide range of physical and psychological symptoms, which detract from their quality of life and often require hospital admission.

Although 52% of deaths from colorectal cancer occur in the over-75 age group, colorectal cancer is nonetheless a significant cause of both premature death and morbidity. The aim of treatment in patients with advanced disease is to improve both the duration and quality of the patient's remaining life.

### Interface A: Project specification and selection

For the specific review, the three drugs were to be evaluated as both monotherapy and combination therapy, in the first- and second-line treatment of patients with advanced colorectal cancer, in relation to: disease progression rates; their relative effect on overall survival and quality-of-life-adjusted survival; their side-effect profiles; their incremental cost-effectiveness in comparison with conventional therapy; and the overall cost associated with the use of these drugs in England and Wales. This NICE TAR was contracted to the team at ScHARR, University of Sheffield.

### Stage 1: Inputs

The NICE TAR Centres work under a contract which involves them conducting a number of TARs each year on a 'call-off contract basis' without a specification of the amount of funding for each one. The average 'cost' of TARs is discussed elsewhere in this report. All TAR teams consist of researchers with experience of conducting health technology assessments. Most of the funding was used for three members of the team who undertook this review and the rest was used for expenses such as inter-library loans, travel and so on. Practitioners with relevant clinical expertise were also brought into the specific team. The expert advisory group provided useful input, especially on clinical questions.

### Stage 2: Processes

For 3 months between October 2000 and January 2001, a systematic review of the literature was conducted involving several databases. The calculations of cost-effectiveness were based on progression-free survival, rather than survival, because when chemotherapy is given subsequent to the allocated first-line regimens, survival cannot be uniquely related to the allocated therapy. The methods proved appropriate at the time and,

according to the appraisal of the review conducted by the Centre for Reviews and Dissemination, the authors' conclusions seemed appropriate in light of the data they presented (CRD, 2003). Nevertheless, the review of the Technology Appraisal, also conducted at ScHARR, adopted a new economic model (NICE, 2005).

### Stage 3: Primary outputs

#### Knowledge production

In relation to clinical effectiveness, the authors of the TAR concluded in the HTA monograph that, "When used as first-line therapy, the combination of either irinotecan or oxaliplatin with an infusional fluorouracil and folinic acid (FU/FA) regimen appears to extend median progression-free survival by 2–3 months compared with FU/FA alone, although with increased toxicity; irinotecan has also been shown to extend overall survival. However, raltitrexed appears to reduce both progression-free and overall survival compared with FU/FA. When used as second-line treatment, irinotecan monotherapy appears to extend median progression-free survival by approximately 1 month and overall survival by approximately 2 months compared with FU/FA alone, again at the cost of increased toxicity. Preliminary data suggest that, as second-line treatment, oxaliplatin plus 5FU may extend median progression-free survival compared with either 5FU or irinotecan monotherapy." (Lloyd Jones and colleagues, 2001, p. vi).

As noted above, progression-free survival was used in place of survival in the economic analysis and this had considerable implications: "Oxaliplatin shows greater improvement than irinotecan in progression-free survival, compared with 5FU, based on our analysis of the progression-free survival curves; however, no survival benefit has been shown in clinical trials with oxaliplatin, whereas it has with irinotecan. For second-line treatment (after which smaller proportions of patients had further chemotherapy compared with after first-line therapy), cost-effectiveness ratios were estimated on the basis of both progression-free survival and survival. The results of the two estimates are different.

"The marginal cost per progression-free year for oxaliplatin compared with the de Gramont 5FU regimen is £23,000. The equivalent cost for irinotecan is £58,400. These figures are obviously dependent on the cost estimates that ... are more uncertain for irinotecan than for oxaliplatin. Second-line treatment with irinotecan (single-agent therapy) is less expensive than the inpatient de Gramont regimen. If it is assumed that all

treatments are given on an outpatient basis, the marginal cost per progression-free year is unchanged for oxaliplatin, £49,000 for irinotecan and £26,400 for second-line irinotecan.

“For second-line treatment, the marginal cost per life-year gained (i.e. based on survival benefit) is zero when irinotecan is compared to inpatient treatment with the de Gramont regimen, £11,180 when compared to outpatient de Gramont, and between £17,700 and £28,200 when compared to BSC [best supportive care] ... . Because there is no benefit in either progression-free survival or survival when treatment with raltitrexed is compared with 5FU, a cost-effectiveness analysis is not appropriate.” (Lloyd Jones and colleagues, 2001, pp. v–vi).

The main publication from the study was the HTA monograph, with no other articles being produced. Writing journal articles from the work of NICE TARs can be problematic when, as in this case, not only has the current position been fully explored in the monograph, but also there are likely to be rapid developments with new studies known to be in the pipeline. In this case, the HTA report has received a considerable number of hits but comparatively few citations. There is the possibility that the number of citations has been under-recorded by the Science Citation Index (SCI). Such an event has been noted in previous circumstances where the lead author’s name has a ‘non-standard’ format, and the figure of six for the current citations was the result of a more thorough exploration of the various possible options than was possible at the time of initial scoring in 2005. The recent large increase in the number of hits during the course of this impact assessment shows that the monograph continues to attract considerable interest:

Lloyd Jones M, Hummel S, Bansback N, Orr B, Seymour M. 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. *Health Technol Assess* 2001;5(25). [at initial project scoring in 2005, 204 despatched, 15,295 hits, 1 cite; current, 205 despatched, 22,671 hits, 6 cites].

#### **Research targeting, capacity building and absorption**

The main researchers have not really been involved in further work in this field, but the ScHARR team were asked to conduct further work in the colorectal field.

#### **Interface B: Dissemination**

Dissemination seemed to be concentrated on the direct links with the NICE Appraisal Committee.

As usual the TAR was considered by the NICE Appraisal Committee when it developed its advice on this topic and team members attended the relevant meeting. They were consulted about various matters, as noted in the minutes of the meeting held on 7 February 2001 (NICE, 2001).

#### **Stage 4: Secondary outputs: informing policy**

NICE Technology Appraisal Guidance No. 33 would appear to have reflected the findings of the TAR when it stated that on balance of clinical and cost-effectiveness, neither irinotecan nor oxaliplatin in combination with 5FU/FA were recommended for routine first-line therapy for advanced colorectal cancer. There were circumstances in which their use was recommended, but raltitrexed was not recommended at all and its use confined to appropriately designed clinical trials (NICE, 2002). The TAR is discussed in the Guidance and clearly informed the debate and even though the cost-effectiveness of irinotecan was reassessed, the Appraisal Committee concluded, “the additional benefits of treatment were not seen to justify the increased cost” (para. 4.3.7). There was considerable criticism of the Guidance, including an editorial in the *British Journal of Cancer* (Saunders and Valle, 2002) claiming that the Guidance was far too restrictive. It seems reasonable to suggest that the TAR was not only considered but played some role in the committee’s decision: the Guidance was broadly in line with the findings in the TAR despite alternative views being expressed.

#### **Stage 5: Applications by practitioners and public**

Although the research team themselves do not seem to have played any role in encouraging practitioners to adopt their findings, the implementation of NICE Guidance is an indirect way of the findings being adopted when the TAR informed the Guidance. The Abacus analysis of the impact of this Guidance suffers from being based on a database with very small numbers. Nevertheless, the figures do indicate that the Guidance was having an impact, with the Abacus report concluding: “This disease area is a very good example of how NICE recommendations can shape prescribing behaviour” (Abacus International, 2005, p. 15).

#### **Stage 6: Impacts or final outcomes**

It is particularly difficult to make any assessment of final outcomes when the overall recommendations were complex and involved use

of the drugs in some circumstances and not others. Nevertheless, perhaps as most clearly seen with the recommendation that raltitrexed should not be used, the Guidance helped enhance the extent to which resources were used appropriately.

### Comments

In this case study, considerable use was made of reviews of documents, including the HTA report, the NICE Guidance and the Abacus report, and also some other documents identified and accessed via the Internet.

In the literature on research impact, the permeability of the interface between the research and policy systems is seen as an important issue (Buxton and Hanney, 1996). This case study again highlights the potential importance of organisational arrangements that ensure the presence of a 'receptor body' (Hanney and colleagues, 2003; Kogan and colleagues, 2006) to receive and use the findings from health research. As in this case, such arrangements can sometimes mean the research can make a considerable impact even though wider dissemination activities are limited.

### References

- Abacus International. *NICE Guidance implementation tracking: data sources, methodology, and results*. Abacus International; 2005. URL: <http://www.nice.org.uk/page.aspx?o=209966>
- Buxton M, Hanney S. How can payback from health services research be assessed? *J Health Serv Res Policy* 1996;1:35–43.
- Centre for Reviews and Dissemination. *Document 277712*. York: CRD, University of York; 2003. URL: <http://144.32.150.197/scripts/WEBC.EXE/nhscred/expand?saan=0000277712>
- Hanney S, Gonzalez-Block M, Buxton M, Kogan M. The utilisation of health research in policy-making: Concepts, examples and methods of assessment. *Health Res Policy Syst* 2003;1:2. URL: <http://www.health-policy-systems.com/content/1/1/2>
- Kogan M, Henkel M, Hanney S. *Government and research: thirty years of evolution*. 2nd ed. Dordrecht: Springer; 2006.
- Lloyd Jones M, Hummel S, Bansback N, Orr B, Seymour M. 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. *Health Technol Assess* 2001;5(25).
- National Institute for Clinical Excellence. *Minutes of the Appraisals Committee Meeting on 7th February 2001*. London: NICE; 2001. URL: <http://www.nice.org.uk/page.aspx?o=15823>

National Institute for Clinical Excellence. *Guidance on the use of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer*. Technology Appraisal Guidance No. 33. London: NICE; 2002.

National Institute for Health and Clinical Excellence. *Irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer*. Technology Appraisal Guidance No. 93. London: NICE; 2005.

Saunders MP, Valle JW. Why hasn't the National Institute been 'NICE' to patients with colorectal cancer? (Editorial). *Br J Cancer* 2002;86:1667–9.

## Case study 16: Effectiveness and cost-effectiveness of imatinib for first-line treatment of chronic myeloid leukaemia in chronic phase: a systematic review and economic analysis (NICE TAR 02/18/01)

### Stage 0: Needs assessment

In this example the topic was really identified for an earlier NICE TAR (Garside and colleagues, 2002) which considered the effectiveness and cost-effectiveness of imatinib for the treatment of chronic myeloid leukaemia (CML), a rare blood cancer with an incidence of 1.0 per 100,000 for men and 0.8 per 100,000 for women. There are three identifiable phases of CML: chronic, accelerated and blast phase, with blast phase being fatal within 3–6 months. CML is not currently curable with conventional chemotherapy or immunotherapy. Patients diagnosed in the chronic phase may expect a median of 3–5 years' survival. Bone marrow transplantation (BMT) offers a cure but is only available to a minority of people. Existing drug treatments included interferon-alpha (IFN- $\alpha$ ) and hydroxyurea. Imatinib is a new treatment that works by blocking the ATP binding site on the BCR-ABL tyrosine kinase. In a NICE Guidance of October 2002 (Technology Appraisal 50), imatinib had already been recommended for the treatment of patients in all phases of the disease who have failed treatment with IFN- $\alpha$  (NICE, 2002). There soon became an issue as to whether the use of imatinib should be expanded so that it became the first-line treatment.

### Interface A: Project specification and selection

NICE set the scope or objective of the assessment as being to evaluate the effectiveness of imatinib as first-line treatment for those with CML in chronic phase compared with IFN- $\alpha$ , hydroxyurea and BMT, and the cost-effectiveness of imatinib

compared with IFN- $\alpha$  and hydroxyurea. Having conducted the initial systematic review, the Peninsular Technology Assessment Group (PenTAG) based at the University of Exeter, was asked to conduct this TAR in conjunction with Southampton Health Technology Assessment, Wessex Institute for Health Research and Development.

### Stage 1: Inputs

The NICE TAR Centres work under a contract which involves them conducting a number of TARs each year on a 'call-off contract basis' without a specification of the amount of funding for each one. The average 'cost' of TARs is discussed elsewhere in this review. All TAR teams consist of researchers with experience of conducting HTAs. The expert advisory group provided useful input, especially on clinical questions.

### Stage 2: Processes

The study was conducted over 4 months in 2002–3. A systematic review of the literature was undertaken. All studies of imatinib were included, along with RCTs of IFN- $\alpha$  compared with hydroxyurea and comparative studies of BMT compared with IFN- $\alpha$ . The assessment included all patient relevant outcome measures reported by the studies, with survival the key outcome measure and the relationship between cytogenetic (bone marrow) response (CR) and survival was considered to be sufficiently strong to support the use of CR as a surrogate outcome measure.

A search of the economic literature revealed no published cost-effectiveness studies comparing imatinib and IFN- $\alpha$ . An independent Markov model was constructed and this was compared with models submitted to NICE by the manufacturer of imatinib, Novartis. There was some iteration with the NICE Appraisal Committee (i.e. the primary customer and user of the report) because, following its first meeting, the authors were requested to review the impact of altering some of the assumptions in the economic model.

### Stage 3: Primary outputs

#### Knowledge production

The key knowledge produced was contained in the TAR presented to the Appraisal Committee and subsequently contained in the HTA monograph. In the latter, the authors state: "Intention-to-treat analysis showed that imatinib was associated with complete CR at 12 months follow-up of 68% compared with 20% for the IFN- $\alpha$  plus Ara-C group ( $p < 0.001$ ). The estimated proportion of

people taking imatinib who had not progressed to accelerated or blast phases at 12 months was 98.5% and 93.1% for IFN- $\alpha$  plus Ara-C ( $p < 0.001$ ) ... Quality of life was better in the imatinib group than the IFN- $\alpha$  group when assessed at 1, 3 and 6 months using the Functional Assessment of Cancer Therapy – Biological Response Modifier instrument ...

"The incremental cost-effectiveness ratio (ICER) of imatinib compared with IFN- $\alpha$  from the independent model was £26,180 per quality-adjusted life-year (QALY) gained (ranging from £13,555 to £51,870) and was relatively robust when subjected to a number of sensitivity analyses. This figure is similar to industry estimates of between £18,000 and £26,000. Imatinib was less cost-effective than hydroxyurea with an ICER of £86,934. Probabilistic analysis showed that if the decision-maker was willing to pay £27,000 per QALY, then imatinib had a greater probability of being cost-effective than IFN- $\alpha$ ." The authors concluded, "Imatinib appears to be more effective than current standard drug treatments in terms of cytogenetic response and progression-free survival, with fewer side-effects." (Dalziel and colleagues, 2004, pp. xi–xii).

One article has been published in addition to the HTA report, but this was published after the completion of the questionnaire.

Dalziel K, Round A, Stein K, Garside R, Price A. Effectiveness and cost-effectiveness of imatinib for first-line treatment of chronic myeloid leukaemia in chronic phase: a systematic review and economic analysis. *Health Technol Assess* 2004;8(28). [at initial scoring of the project in 2005, 57 despatched, 8,684 hits, 1 cite; current, 112 despatched, 19,718 hits, 2 cites].

Dalziel K, Round A, Stein K, Garside R. Cost-effectiveness analysis of imatinib compared to interferon alpha and hydroxurea for first-line treatment of chronic myeloid leukaemia. *Pharmacoeconomics* 2005;23:515–26. [2004 journal impact factor, 2; no cites].

#### Research targeting, capacity building and absorption

There was no real targeting of further research on this topic by the team, but they did set out an agenda for further research on this use of imatinib. The main researcher has used the techniques developed for the modelling in a subsequent model. It was suggested that because the imatinib TAR was seen as a successful study, it possibly played some part in the gaining of an expanded contract for PenTAG as one of the seven TAR centres providing NICE TARs.

The study did not contribute to any higher degrees, but one of the team developed modelling skills.

### **Interface B: dissemination**

A couple of presentations were given to academic audiences, but more importantly the whole structure of the NICE appraisal process gives the researchers a direct feed into the NICE appraisal policy-making process. Indeed, not only was the TAR submitted to the Appraisal Committee, but also team members were invited to provide further analysis during a meeting of the committee.

Generally, the haematologists who would be the main clinical users of the drug were keen to be allowed to prescribe it and, therefore, there was little need to give presentations to them about the research.

### **Stage 4: Secondary outputs: informing policy**

It is clear from the NICE Technology Appraisal (No. 70), and the minutes of the Appraisal Committee meeting, that the evidence from the TAR played an important part in the considerations and development of the Appraisal. The guidance stated: "imatinib is recommended as first-line treatment for people with Philadelphia-chromosome-positive chronic myeloid leukaemia (CML) in the chronic phase." (NICE, 2003, p. 4). This is in line with the conclusions of the TAR, as were some of the more detailed points including the recommendations for further research. This guidance went further than the previous one (NICE, 2002), which recommended the use of imatinib only as a treatment option in those who have failed treatment with IFN- $\alpha$ .

The two NICE decisions to recommend imatinib have been subject to considerable debate in various discussions about the role of NICE (references including Rawlins and Culyer, 2004 and Raftery, 2006) and about the appropriate comparator in this Appraisal.

### **Stage 5: Applications by practitioners and public**

Discussion of the impact of the NICE TAR on practitioners and patients cannot be separated from analysis of the role of NICE. There was considerable support for use of the drug, but it was not widely available on the NHS until NICE reported on it. Therefore, to the extent that the TAR seems to have influenced the positive recommendation for the use of imatinib, the research could be viewed as being likely to have influenced practice in an important way. In terms

of its take-up by clinicians following the NICE recommendations, however, the TAR would probably have little direct impact because clinicians already wanted to adopt it. One possible area where the publications from the study might potentially have some impact in raising awareness of the drug is in relation to the non-specialists to whom the patients might first present and, as noted above, the electronic version of the HTA report has been accessed fairly frequently.

### **Stage 6: Impacts or final outcomes**

The report described the benefits from the use of imatinib but pointed out that only 18-month follow-up data were available. Both the TAR and the Guidance recommended that data be collected about the long-term effectiveness. A clear health gain in the short term results from its use, in terms of both length and quality of life, and the potential numbers of users is known reasonably well. There are about 2660 patients, mostly in the chronic phase, of whom about 2500 would be eligible for use. According to the Abacus report, about 500 patients were receiving the treatment about 1 year after the initial Guidance and therefore at about the time the NICE Guidance 70 was issued (Abacus International, 2005). The incidence of CML is about 600 cases per year. A full analysis of the health gain from the use of imatinib will have to wait until not only are the take-up figures known, but also the long-term effectiveness is assessed. The drug is at about the limits of the costs for positive recommendations from NICE and so there are no cost savings resulting from its use.

### **Comments**

This study illustrates that a series of sources can each contribute to building a picture about the impact of a study. The questionnaire and documentary analysis of the HTA report can provide a useful start, the interview takes things further but then review of further documents, including the NICE Guidance and minutes of NICE meetings, is also useful. Then sometimes documentary sources can go even further, as with the various discussions about the Guidance in this case.

In the literature on research impact, the permeability of the interface between the research and policy systems is seen as an important issue (Buxton and Hanney, 1996). This case study again highlights the potential importance of organisational arrangements that ensure the presence of a 'receptor body' (Hanney et al., 2003; Kogan and colleagues, 2006) to receive and use the findings from health research.

## References

Abacus International. *NICE Guidance implementation tracking: data sources, methodology, and results*. Abacus International; 2005. URL: <http://www.nice.org.uk/page.aspx?o=209966>

Buxton M, Hanney S. How can payback from health services research be assessed? *J Health Serv Res Policy* 1996;**1**:35–43.

Dalziel K, Round A, Stein K, Garside R, Price A. Effectiveness and cost-effectiveness of imatinib for first-line treatment of chronic myeloid leukaemia in chronic phase: a systematic review and economic analysis. *Health Technol Assess* 2004;**8**(28).

Garside R, Round A, Dalziel K, Stein K, Royle R. The effectiveness and cost-effectiveness of imatinib in chronic myeloid leukaemia phase: a systematic review. *Health Technol Assess* 2002;**6**(33).

Hanney S, Gonzalez-Block M, Buxton M, Kogan M. The utilisation of health research in policy-making: concepts,

examples and methods of assessment. *Health Res Policy Syst* 2003;**1**:2. URL: <http://www.health-policy-systems.com/content/1/1/2>

Kogan M, Henkel M, Hanney S. *Government and research: thirty years of evolution*. 2nd ed. Dordrecht: Springer; 2006.

National Institute for Clinical Excellence. *Guidance on the use of imatinib for chronic myeloid leukaemia*. Technology Appraisal Guide No. 50. London: NICE; 2002.

National Institute for Clinical Excellence. *Guidance on the use of imatinib for chronic myeloid leukaemia*. Technology Appraisal Guide No. 70. London: NICE; 2003.

Raftery J. Review of NICE's recommendations, 1999–2005. *BMJ* 2006;**333**:1266–8.

Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. *BMJ* 2004;**329**:224–7.





## Appendix 9

# Scoring and re-scoring of the case study projects

### Aim of scoring payback on the basis of knowledge gathered prior to case studies

Those HTA projects selected for case studies were scored using the information available prior to case studies and then re-scored on the basis of the greater data available from the case studies. This helped assess the appropriateness of using questionnaires and thus fed into the discussion of recommendations for a regular monitoring system to be developed by the HTA Programme.

### The payback categories on which scoring was based

The payback scoring prior to use of the data collected in the case studies concentrated on data in the following categories from the Buxton and Hanney payback framework (see Chapter 3): knowledge production; research benefits; impact on policy making; and impact on behaviour. A score for impact on behaviour/practice is not identical with a score for health gain but is an intermediary approach (it is linked to the penultimate stage in the payback logic model rather than the final stage of health and economic benefits). To make things more manageable for scoring, we broke each category down into two items. This meant that in total there were eight scoring scales.

In the actual case studies, we aimed to conduct analysis relevant for all payback categories wherever possible with the hope that this would show up as something that the case studies could do that questionnaires could not. The intention was that the rescoring of the case studies could, therefore, cover all payback categories, but insufficient data were gathered to make this practical.

### Number of scoring levels

For each of the eight scoring scales, the levels used ranged from 0 to 5, with frequent use of a score in

between, at the 4.5, 3.5, 2.5, 1.5 or 0.5 level. The exemplar descriptors given below (*Boxes 7–10*) for the main levels for each of the eight scales indicate the type of impact that could be expected at each level: clearly there had to be room for flexible interpretation of the diverse situations that arose.

### Devising the scoring scales and levels

The scales were generally devised to reflect what was thought appropriate for HTA Projects. The top levels for the scales for Knowledge Production and Research Capacity Building and Targeting (see below) reflected what was considered to be achieved by the best of the projects from the HTA portfolio overall, irrespective of whether they were included in the set of case studies. Thus for the HTA reports the top level was: “Over 750 copies of the HTA report despatched, the electronic version of the report was hit over 60,000 times, and been cited 75 times. (For the maximum score all are required.)” This was based on the position in Autumn 2005 of HTA Project 96/17/01, *Systematic review of treatments for atopic eczema*, of which Hywel Williams was the Principal Investigator: 774 copies despatched, 81,689 hits and 104 citations. There are sometimes reservations about what exactly a count of the number of hits means in terms of impact, because some reports with many hits are known to appear on methodological reading lists of courses for health professionals. Nevertheless, both the study by Williams, and the one with the most hits (108,032 in Autumn 2005 and 150,841 in April 2006) were substantial studies.

The scales for impact on policy and practice were also partly informed by the earlier work from HERG.

The eight scales are presented in *Boxes 7–10*. Each box contains the two scales for the relevant payback category: knowledge production; research benefits; impact on policy making; and impact on behaviour.

**BOX 7** *Scoring scales for Knowledge Production***a Scale for Knowledge Production: HTA report**

- 5 – Over 750 copies of the HTA report despatched, the electronic version of the report was hit over 60,000 times, and been cited 75 times. (For the maximum score all are required.)
- 4 – Over 500 copies of the HTA report despatched, hits over 30,000, citations over 20.
- 3 – Over 250 copies of the HTA report despatched, hits over 20,000, citations over 15.
- 2 – Over 100 copies of the HTA report despatched, hits over 5000, citations over 5.
- 1 – Over 50 copies of the HTA report despatched, hits over 2000, 1 citation.
- 0 – No HTA report.

**b Scale for Knowledge Production: publications**

- 5 – The project produced at least eight publications, most of which were peer-reviewed articles, and appeared in forms that were likely to make an impact on the intended audience, e.g. appearing in high impact factor journals. At least one publication cited more than 50 times.
- 4 – The project produced at least five peer-reviewed publications (not abstracts or conference proceedings), including one that has been cited more than 25 times.
- 3 – The project produced at least three peer-reviewed publications (not abstracts or conference proceedings).
- 2 – The project produced at least one peer-reviewed publication or one highly relevant for the target audience.
- 1 – The project produced internal but no external publications.
- 0 – The project produced no publications or internal report.

**BOX 8** *Scoring scales for Research Benefits***a Scale for Research Benefits: capacity building**

- 5 – The project made a considerable contribution to at least two research/higher degrees, such as PhDs/MDs.
- 4 – The project either made a considerable contribution to at least one research/higher degree, or a moderate contribution to at least two.
- 3 – The project made a moderate contribution to at least one research degree.
- 2 – The project made some contribution to at least one research degree.
- 1 – The project made some contribution to research capacity building in some other way.
- 0 – The project made no contribution to research degrees or research capacity building in any other way.

**b Scale for Research Benefits: research targeting**

- 5 – The project made a considerable contribution to more than one follow-on project by the team and/or by others and the importance of these projects should be indicated by, for example, being of at least double the value of the original project.
- 4 – The project made a contribution to more than one follow-on project, considerable in at least one case, and the importance of these projects should be indicated by, for example, being of at least the value of the original HTA project.
- 3 – The project made a contribution to more than one follow-on project, moderate in at least one case.
- 2 – The project made a moderate contribution to one follow-on project, or any contribution to more than one follow-on project.
- 1 – The project made a contribution to at least one follow-on project.
- 0 – The project made no contribution to targeting of future research.

**Sources of data used for applying the scoring scales to the projects**

Various sources of information were available for each of the 16 projects and could be used in applying the scoring scales. The first two bullet points below describe the information available for each project and how it was to be used. Then a final bullet point gives some additional information that was used for a few projects when scoring questions 3a (and 3b):

- First, for each of the 16 case study projects, SH prepared a separate list of the main publications which combined information on the HTA report, journal articles, editorials and sometimes letters and book chapters. Cochrane Reviews were counted as peer-reviewed articles. For the HTA report, the list contained not only the number of citations, if any, received by the time of scoring, but also the number of copies despatched and the number of times the web version of the report had been hit. For each

**BOX 9** Scoring scales for Informing Policy making**a Scoring scale for Informing Policy making: the nature of the policy**

- 5 – The project made an impact on a substantial policy of an international body or substantial policies of at least two national governments.
- 4 – The project made an impact on at least one policy from a national policy-making body such as NICE.
- 3 – The project made an impact on at least one policy from a national professional body.
- 2 – The project made an impact on the policy making of at least one local unit of the health service.
- 1 – A claim for impact was made but no details given, or details given of a claim for expected future impacts.
- 0 – The project made no impact on policies.

**b Scoring scale for Informing Policy making: degree of impact**

- 5 – The policy was almost solely based on the project's evidence in a direct instrumental way.
- 4 – The project made a considerable impact on the policy.
- 3 – The project made a moderate impact on the policy in an instrumental way, or made an important contribution at a conceptual level to the policy discussions.
- 2 – The project made some identifiable impact on the policy.
- 1 – A claim for impact was made but no details given, or details given of a claim for expected future impacts.
- 0 – No impact on policy making.

**BOX 10** Scoring scales for Informing Behaviour**a Scoring scale for Informing Behaviour: the level of impact**

- 5 – The project made an impact on behaviour in more than one country.
- 4 – The project made an impact on behaviour at a national level.
- 3 – The project made an impact on the behaviour of at least one team of practitioners or managers, or at least one group of patients/members of the wider public.
- 2 – The project made an impact on behaviour of at least one or more practitioner, manager, patient or member of the public.
- 1 – A claim for impact made but no details given, or details given of a claim for expected future impacts.
- 0 – The project made no impact on behaviour.

**b Scoring scale for Informing Behaviour: the degree of impact**

- 5 – The changed behaviour was almost solely based on the project's evidence in a direct instrumental way.
- 4 – The project made a considerable impact on the behaviour.
- 3 – The project made a moderate impact on the behaviour.
- 2 – The project made some identifiable impact on the behaviour.
- 1 – A claim for impact made but no details given, or details given of a claim for expected future impacts.
- 0 – No impact on behaviour.

journal article, the list contained details about the number of citations and the journal impact factor. This project list provided the information required to complete the first question on the scoring sheet, 1a. Scale for Knowledge Production: HTA report, and in most cases the second question, 1b. Scale for Knowledge Production: Publications. In a few cases the list of publications did not contain any publications in the above categories, but some information was available about publications such as abstracts and presentations from the original list of publications for each project held by NCCHTA and from the data sheet described next.

- Second, there were the main data sheets produced for each project by DC. Each one included all the data from the questionnaires, plus some data from the files of the NCCHTA.

The datasheets contained information relevant for scoring the remaining questions (2a–4b). Also in Section E of the datasheet there was information about published abstracts and presentations not included on the lists described above.

- Finally, there was some information available from lists produced by the NCCHTA about the HTA reports cited in documents such as NICE Guidance and reports from the NSC (see Chapter 4). All three NICE TARs in our list of projects (case studies 14, 15, 16) were included on the list of NICE TARs that related to NICE Guidance that had been produced. This meant that irrespective of what was said on the questionnaire we assumed they were used in policymaking at a national level (scale 3a) and such reports usually play an important part in

the considerations of the NICE Appraisal Committee (relevant for scale 3b) because they are specifically commissioned to inform the Committee's deliberations as it formulates the Guidance. On the list supplied of HTA reports that informed the national policy making by the NSC, there is an account of the part played by the report included in case study 10: "On the basis of this report, the NSC has recommended a systematic approach to case finding for hypercholesterolaemia in preference to whole population screening." Again, such evidence was used for scales 3a and 3 b.

## Conducting the scoring and re-scoring

Using the data available before the case studies were completed, the 16 projects were scored by SH and Teresa Jones (TJ), a researcher from HERG who was not otherwise involved in the project. [Due to an administrative oversight, SH did not score three of the eventual 16 using the data originally available until after the case study had been conducted.]

Both SH and TJ re-scored the projects based on the full data available from the drafts of the case studies. Where, however, there had been increases in the number of copies of reports despatched, hits and citations the additions were not taken into account in the re-scoring and neither were additional publications that came after the date of the original scoring. This is because these were due to the passage of time and not to different methods of data collection. Where, however, the case study revealed additional items that had been published prior to the date of the initial scoring, but which were not identified until the case study, these were included in the information that was used in the re-scoring. The re-scoring was undertaken after the first draft of the case studies had been completed, rather than waiting for the final versions, for various reasons. These included the problems being caused by increasing amounts of additional data available that were coming in as a result of the passage of time rather than just the different methods of data collection.

## Assessing inter-rater reliability

It was thought useful to assess inter-rater reliability and a weighted kappa approach was used. The kappa statistic is scaled where a value 0 denotes agreement that would be observed by chance and

1 denotes perfect agreement. To interpret the kappa statistic, six terms are used: poor (which is a score of less than zero); slight (0.00–0.20); fair (0.21–0.40); moderate (0.41–0.60); substantial (0.61–0.80) and almost perfect (0.81–0.99).

The results of the kappa analysis are also shown in *Table 17*, with the scores from the two scorers ( $x$  and  $y$ ) shown for each of the eight scales for each of the 16 projects. In addition to the kappa analysis, the table also shows the actual scores given in the scoring and re-scoring, but the order in which the 16 projects are listed has been randomised.

The results for the two sets of initial scores for each of the eight scales (and for a total score for each project that, just for this exercise, was a sum of the eight scores for the case study) range from 0.35 to 0.88, with all except two in the substantial or almost perfect categories. As might be expected, the highest scores come from those scales where the most mechanistic approach was possible, that is, the scales for the HTA reports, capacity building and research targeting. The lowest came from those that were most subjective and where the data gathered in the questionnaire were least fully completed, that is, degree of impact on policy and practice. The kappa scores for the 16 projects range from 0.29 to 1.

[Before CS left the project, she had scored 15 of the projects originally selected for case studies that SH had also scored. Subsequently the Principal Investigators from some of those projects declined to participate in case studies and further projects had to be selected. These were scored by SH but not CS. Nevertheless, it was thought useful to assess inter-rater reliability at a preliminary stage. Therefore, just for testing inter-rater reliability, for each of the 15 projects scored by both CS and SH the combined scores from all eight scales were considered, using the weighted kappa approach. The degree of inter-rater reliability for these 15 scores was found to be 0.73 (substantial), which was very similar to the later score of 0.69 for the equivalent scoring between SH and TJ.]

Kappa analysis was also conducted on the re-scores based on the data from the case studies and this is also shown in *Table 17*. The kappa analysis for the re-scoring again showed generally high levels of agreement. Although there was some decrease in the previously very high agreement levels for research targeting and the nature of the policy, there was some increase in the previously low level of agreement for degree of impact on

policy making. Both of these shifts might have been expected. In the case of research targeting and the nature of the policy making there were often already clear, albeit brief, questionnaire responses and the increased information from the survey provided greater detail but also additional information that could be variously interpreted. By contrast, the data available prior to the case studies about the degree of impact on policy making were severely limited and the additional information provided a somewhat clearer picture. [It became clear that for three projects the re-scores given for the two items related to impact on research had become transposed and these were therefore corrected and the kappa analysis re-run to give the figures provided here.]

## Comparing the questionnaires and case studies

The kappa scores for inter-rater reliability were considered to be sufficiently satisfactory to make it reasonable to proceed with an analysis comparing the initial scoring and the re-scoring. This was undertaken to address the issue of whether the survey respondents seemed to be exaggerating the impact of their projects. It is difficult to make strong comparisons because the case studies again relied heavily on the information from the Principal Investigator and, therefore, there is considerable scope for bias to be built into both questionnaires and case studies. Nevertheless, attempts were made in the case studies to verify claims made by conducting some documentary analysis (see Chapter 6).

A detailed comparison was made of the level of agreement between the two scorers on the basis of the direction, not amount, of change between their scoring and re-scoring. The results are shown in *Table 18* and reveal that there was a considerable agreement between the two scorers. There were 16 projects scored on nine items (eight scales and an overall score), giving a total of 144 cells. Of these the two scorers ( $x$  and  $y$ ) agreed for 95 cells (either both stayed the same or both moved in the same direction). In those instances where both scorers had changed their scores there were 40 examples of both giving a higher score in the second round of scoring (i.e. based on the case studies) than in the first round (i.e. based on the questionnaires) and only five where both scorers gave a lower score.

Of the 49 cells where the two scorers disagreed, there were only 13 where they went in opposite

directions; more usually, one changed their score between the two rounds and the other remained the same. In 30 out of the 36 instances where one scorer was higher in one round and the other the same, it was a case of the one scorer giving a higher score in the second round.

This analysis suggests that, in general, the further data collection involved in the case studies indicated that researchers were not over-stating the impact of their projects in their questionnaires. This picture was further strengthened by the additional data gathered in the case studies between the first and final drafts. The implications of this are discussed below, but first we consider some lessons learnt about the conduct of the studies.

## Lessons learnt about scoring

Various comments about the questionnaire made by respondents were reported in Chapter 5. Analysis of the scoring and re-scoring raised further questions and issues relevant for the questionnaire and also the scoring based on the information gathered through the questionnaire and the case studies. Some of the main questions raised by the current study are listed below.

### Issues linked to the questionnaire

1. In the questionnaire and database, better definitions of publications are required to deal with issues such as how far published abstracts should be counted, especially if there was also a subsequent article based on the abstract.
2. In some of the case studies, it became clear that there was some contribution to career development, but this was not asked about specifically in the questionnaire.
3. In the questionnaire, there were separate sections for the level of impact on behaviour and the degree to which the research had caused the impact, but the respective questions in the policy section of the questionnaire were not separated out. This needs to be regularised with a greater attempt to capture information relevant for scoring.

### Issues linked to scoring the data collected about capacity development

4. Consideration is needed on how best to score career development as part of the score for capacity development.
5. Clarification is needed as to whether all higher/research degrees should have equal status.

TABLE 17 Inter-rater reliability of the initial scoring and re-scoring of the 16 projects

Project <sup>a</sup>	1. Knowledge production				2. Research benefits				3. Information policy making				4. Changed behaviour				Total score (sum)	Kappa score		
	a. HTA report		b. Publications		a. Capacity building		b. Research targeting		a. Nature of the policy		b. Degree of impact		a. Level of impact		b. Degree of impact					
	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y				
A	3	3.5	2	2.5	4	4	4	4.5	4	4	5	4	4	4	4	4	1	30	27.5	0.46
B	2	2.5	2	0	0	0	0	0	4	4	4	1	0.5	1	0.5	1	0.5	14	8.5	0.58
C	1	1	4.5	4.5	5	4.5	3	3.5	4	4	3	3	4	3	4	3	4	24.5	28.5	0.63
D	0.5	0.5	2	2.5	0	0	2.5	2.5	3	4	2	1	1	1	1	1	2	12	13.5	0.77
E	0	0	2	2.5	0	0	2.5	4	1	0.5	1	0.5	1	1	1	1	1	8.5	9.5	0.75
F	1	1.5	2	2.5	0	0	1	1	1	4	1	3	2	2.5	4	5	5	12	19.5	0.53
G	1	1.5	3.5	3.5	0	0	3	3	0	0	0	0	0	0	0	0	0	7.5	8	0.89
H	2	2	2.5	2.5	0	0	3.5	3.5	4	4	2	3.5	4	2	1	4	4	19	21.5	0.50
I	1	1.5	2	0	0	0	2	2	4	4	4	3	4	4	5	3	3	22	17.5	0.66
J	2	2.5	2	2	2	2	1	1.5	4	4	4	2.5	4	4	2	3	3	21	21.5	0.63
K	0	0	1.5	1.5	0	0	0	0	1	0.5	1	0.5	0	0	0	0	0	3.5	2.5	0.79
L	0.5	0.5	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2.5	2.5	1.00
M	0.5	0.5	2	1.5	3	3	1	0.5	0.5	0	0.5	0	1	2	1	2	2	9.5	9.5	0.60
N	3	3.5	5	4	3	1	2	3	4	4	1	4	1	4	1	3	3	20	26.5	0.29
O	3	3	5	5	0	0	5	5	4.5	4.5	4	2	1	1	1	1	0.5	23.5	21	0.86
P	0	0	1.5	1.5	0	0	0	0	1	1	1	1	0.5	1	0.5	1	0.5	5.5	4.5	0.79
Kappa	0.83		0.68		0.88		0.85		0.79		0.35		0.63		0.46		0.69			
<b>(b) Second scoring method using case study notes and questionnaires</b>																				
Project <sup>a</sup>	1. Knowledge production				2. Research benefits				3. Information policy making				4. Changed behaviour				Total score (sum)	Kappa score		
	a. HTA report		b. Publications		a. Capacity building		b. Research targeting		a. Nature of the policy		b. Degree of impact		a. Level of impact		b. Degree of impact					
	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y				
A	3	3.5	3.5	3.5	4	4	3.5	4	5	4.5	2	5	1	1	1	2	23	27.5	0.59	
B	2	2.5	2	0	0	0	1	0	4	4	3	3.5	4	4	2	3.5	18	17.5	0.67	
C	1	1	4.5	4.5	5	5	3	3.5	3	4.5	2	4	4	4	2	3.5	24.5	30	0.60	
D	0.5	0.5	3	3	0	0	3	3.5	3	4	2	3	4	2	2	2	17.5	18	0.58	
E	0	0	1.5	2.5	1	1	3	4	1	0.5	1	0.5	2	1	1	1	11.5	10.5	0.62	
F	1.5	1.5	2.5	2.5	1	1	0.5	2	4	4	0	1.5	3	3	5	4.5	13.5	20	0.57	

continued

TABLE 17 Inter-rater reliability of the initial scoring and re-scoring of the 16 projects (cont'd)

Project <sup>a</sup>	1. Knowledge production				2. Research benefits				3. Information policy making				4. Changed behaviour				Total score (sum)	Kappa score						
	a. HTA report		b. Publications		a. Capacity building		b. Research targeting		a. Nature of the policy		b. Degree of impact		a. Level of impact		b. Degree of impact									
	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y								
G	1	1.5	3.5	3.5	2	1.5	3	3.5	4	4	2	1.5	0	0.5	0	0.5	15.5	16.5	0.57					
H	2	2	3	3	0	0	3	4	2	3	4	4	3	3	4	4	21	23	0.75					
I	1	1.5	2	0	1	1	2	1	4	4	4	4	4	4	4	3	22	18.5	0.61					
J	2	2.5	2	2	2	2	2	2	4	4	4	3	4	4	2	3.5	22	23	0.66					
K	0	0	0	2	0	0.5	0	0	1	1	1	1	1	1	1	1	4	6.5	0.64					
L	0.5	0.5	2.5	2	4	3	0	0	0	0	0	0	0	0	0	0	7	5.5	0.85					
M	0.5	0.5	2	2	3	3	2	2	1	1	1	1	1	2	1	2	11.5	13.5	0.76					
N	4	3.5	5	4.5	2	1.5	3	3.5	4	4	2	3.5	4	3.5	1	4	25	28	0.44					
O	3	3	5	5	1	1.5	5	5	4	4.5	5	4	1	4	1	1	25	28	0.65					
P	0	0	2	2	1	0	0	0	1	1	1	1	1	0.5	1	0.5	7	5	0.65					
Kappa	0.86		0.73		0.83		0.66		0.65		0.58		0.68		0.51		0.68							
<b>Kappa values Agreement</b>	Below 0.0 Poor				0.00-0.20 Slight				0.21-0.40 Fair				0.41-0.60 Moderate				0.61-0.80 Substantial				0.81-1.00 Almost perfect			

x and y indicate the scores for the two different raters.

<sup>a</sup> The project numbers were randomised and re-ordered.

TABLE 18 Comparison of scoring methods: questionnaires only versus case studies and questionnaires

Scoring	1. Knowledge production		2. Research benefits		3. Informing policy making		4. Changed behaviour		Total project scores for the eight scales (without sum)	Total project scores for the eight scales with sum
	a. HTA report	b. Publications	a. Capacity building	b. Research targeting	a. Nature of the policy	b. Degree of impact	a. Level of impact	b. Degree of impact		
<b>Scores where x and y agree</b>										
Method 1 = method 2	14	6	6	5	8	3	5	3	0	50
Method 1 > method 2	0	0	0	1	1	1	1	1	0	5
Method 1 < method 2	0	4	6	4	3	4	4	2	13	27
Total	14	10	12	10	12	8	10	6	13	82
<b>Scores where x and y disagree</b>										
Method 1 = 2 for 1 rater, method 1 > 2 for other rater	0	1	0	1	2	0	0	1	1	5
Method 1 = 2 for 1 rater, method 1 < 2 for other rater	2	4	3	3	1	4	4	7	2	28
Method 1 > 2 for 1 rater, method 1 < 2 for other rater	0	1	1	2	1	4	2	2	0	13
Total	2	6	4	6	4	8	6	10	3	46
Total number of projects scored	16	16	16	16	16	16	16	16	16	128



### Issues linked to scoring the data collected about impact on policy and practice

6. Clarification is needed about what to do in cases where the follow-on project has had the impact, especially where the follow-on project used some of the data from the original study.
7. Currently the scoring attempts to cover both the level of impact (for example, unit level or national) and the degree of impact on the policy decision, but the scope of the policy is not properly covered, that is, it could be a whole national policy or one small part of a guideline at national level.
8. Clarification is needed as to exactly how negative findings should be counted in relation to making an impact on policy and behaviour if they helped to limit the adoption of a technology that might otherwise have been adopted.
9. Clarification is needed about how to score the impact from a project if it has confirmed current behaviour.
10. NICE TARs raise important points about the scoring for the degree of impact on policy and practice because, although they are automatically considered in the NICE Appraisal, how far they, as opposed to the original research, should be viewed as the key evidence is an important issue. Sometimes similar issues arise in relation to secondary reviews in general.
11. Clarification is needed that a high score should be given for the degree of impact on practice for those projects, especially NICE TARs, that have influenced the policy but for which the impact on practice comes almost entirely through the policy as opposed to any exposure to the original research.

### Dealing with expected impacts

12. Further consideration is needed about how to score claims for expected impacts. This is particularly complicated when considering how to score the degree of impact.

### The relationship between scoring and re-scoring

13. There were difficulties concerned with the re-scoring in that some of the data gathered in case studies (beyond publications) reflect the passage of time, but this was viewed as too complex to attempt to isolate from the greater information gained from using the case study method. The lesson might be that in any future assessment where the use of comparative scoring is planned, the case study

should be conducted closer to the time that the questionnaire was completed.

### Presenting the scoring data

14. Careful consideration would be necessary before scores were presented and used in any way that aggregates the various scores for a project into one total figure. It would be desirable to consult the potential users of any scoring report about the weights to give to different items. For example, in relation to the scores for questions 9a and 9b it would probably be best to present a product of the two, rather than any type of summation, because even if the project has been considered at a national or international policymaking level, if it was not very influential in the policymaking considerations it probably should not receive the highest scores. Similar points apply to 10a and 10b. Furthermore, any move towards an aggregate score is contrary to the multi-dimensional approach previously adopted in the payback framework and the implications of this, especially for researchers, would be controversial. It is likely that at most a type of payback profile should be used.

It is clear that considerable development work would be necessary before the scoring system could be applied on a regular basis.

### Discussion and implications

Several related conclusions can be drawn from the above analysis. The inter-rater reliability between the scorers was reasonably good. This provided some confidence for then using that scoring to conduct some analysis of the methods used in the impact assessment. This analysis suggested that the data being supplied by Principal Investigators in questionnaires do not routinely provide an exaggerated picture of the impact. Particularly if the various suggestions for improving the questionnaires were followed, this might indicate that it would be reasonable to use the questionnaires for regular monitoring of the NHS HTA Programme.

Furthermore, if the scoring itself is seen as reasonable, and account is taken of the points noted in the above analysis of the scoring process, then it might be appropriate to extend the scoring to the whole set of projects for which questionnaires were returned in the current assessment. This might assist further analysis of the

impact arising from the various types of research conducted in the NHS HTA Programme and also help inform any decision about using a scoring process as part of a system of regular monitoring.

The current study was partly intended to identify the most appropriate methods for assessing the impact made by the NHS HTA Programme. This analysis of scoring has made some contribution to

the substantive assessment of the impact by indicating that the level of impact being shown across the programme by the questionnaires might not be an exaggeration. Furthermore, the analysis indicates that further progress in developing a scoring system might be worthwhile as part of an approach to regular monitoring or impact assessment, but many issues still need to be addressed.



# 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 Ukoumunne 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 pre-operative 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, Frementle 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, Reimsma 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, Heny 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, Munafa 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- $\beta$  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<sup>®</sup>) 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  $\alpha$ -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, Griebisch 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, Braunholtz 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 Woloshnowych 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, Sakarovitch 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, Castelnovo 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, Dawney 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, Cogan 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 Castelnovo 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, Griesch 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<sup>®</sup> 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 Castelnuovo 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 non-steroidal 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 WY, 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, McKinlay 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 11, 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 chlamydia 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 Fayter 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). Home-based 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 anticoagulation 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, Düндar 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, Snowdon 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.





# Health Technology Assessment Programme

**Director,**  
**Professor Tom Walley,**  
 Director, NHS HTA Programme,  
 Department of Pharmacology &  
 Therapeutics,  
 University of Liverpool

**Deputy Director,**  
**Professor Jon Nicholl,**  
 Director, Medical Care Research  
 Unit, University of Sheffield,  
 School of Health and Related  
 Research

## Prioritisation Strategy Group

### Members

**Chair,**  
**Professor Tom Walley,**  
 Director, NHS HTA Programme,  
 Department of Pharmacology &  
 Therapeutics,  
 University of Liverpool

Professor Bruce Campbell,  
 Consultant Vascular & General  
 Surgeon, Royal Devon & Exeter  
 Hospital

Professor Robin E Ferner,  
 Consultant Physician and  
 Director, West Midlands Centre  
 for Adverse Drug Reactions,  
 City Hospital NHS Trust,  
 Birmingham

Dr Edmund Jessop, Medical  
 Adviser, National Specialist,  
 Commissioning Advisory Group  
 (NSCAG), Department of  
 Health, London

Professor Jon Nicholl, Director,  
 Medical Care Research Unit,  
 University of Sheffield,  
 School of Health and  
 Related Research

Dr Ron Zimmern, Director,  
 Public Health Genetics Unit,  
 Strangeways Research  
 Laboratories, Cambridge

## HTA Commissioning Board

### Members

**Programme Director,**  
**Professor Tom Walley,**  
 Director, NHS HTA Programme,  
 Department of Pharmacology &  
 Therapeutics,  
 University of Liverpool

**Chair,**  
**Professor Jon Nicholl,**  
 Director, Medical Care Research  
 Unit, University of Sheffield,  
 School of Health and Related  
 Research

**Deputy Chair,**  
**Dr Andrew Farmer,**  
 University Lecturer in General  
 Practice, Department of  
 Primary Health Care,  
 University of Oxford

Dr Jeffrey Aronson,  
 Reader in Clinical  
 Pharmacology, Department of  
 Clinical Pharmacology,  
 Radcliffe Infirmary, Oxford

Professor Deborah Ashby,  
 Professor of Medical Statistics,  
 Department of Environmental  
 and Preventative Medicine,  
 Queen Mary University of  
 London

Professor Ann Bowling,  
 Professor of Health Services  
 Research, Primary Care and  
 Population Studies,  
 University College London

Professor John Cairns,  
 Professor of Health Economics,  
 Public Health Policy,  
 London School of Hygiene  
 and Tropical Medicine,  
 London

Professor Nicky Cullum,  
 Director of Centre for Evidence  
 Based Nursing, Department of  
 Health Sciences, University of  
 York

Professor Jon Deeks,  
 Professor of Health Statistics,  
 University of Birmingham

Professor Jenny Donovan,  
 Professor of Social Medicine,  
 Department of Social Medicine,  
 University of Bristol

Professor Freddie Hamdy,  
 Professor of Urology,  
 University of Sheffield

Professor Allan House,  
 Professor of Liaison Psychiatry,  
 University of Leeds

Professor Sallie Lamb, Director,  
 Warwick Clinical Trials Unit,  
 University of Warwick

Professor Stuart Logan,  
 Director of Health & Social  
 Care Research, The Peninsula  
 Medical School, Universities of  
 Exeter & Plymouth

Professor Miranda Mugford,  
 Professor of Health Economics,  
 University of East Anglia

Dr Linda Patterson,  
 Consultant Physician,  
 Department of Medicine,  
 Burnley General Hospital

Professor Ian Roberts,  
 Professor of Epidemiology &  
 Public Health, Intervention  
 Research Unit, London School  
 of Hygiene and Tropical  
 Medicine

Professor Mark Sculpher,  
 Professor of Health Economics,  
 Centre for Health Economics,  
 Institute for Research in the  
 Social Services,  
 University of York

Professor Kate Thomas,  
 Professor of Complementary  
 and Alternative Medicine,  
 University of Leeds

Professor David John Torgerson,  
 Director of York Trial Unit,  
 Department of Health Sciences,  
 University of York

Professor Hywel Williams,  
 Professor of  
 Dermato-Epidemiology,  
 University of Nottingham

## Diagnostic Technologies & Screening Panel

### Members

#### Chair,

**Dr Ron Zimmern**, Director of the Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge

Ms Norma Armston, Freelance Consumer Advocate, Bolton

Professor Max Bachmann, Professor of Health Care Interfaces, Department of Health Policy and Practice, University of East Anglia

Professor Rudy Bilous, Professor of Clinical Medicine & Consultant Physician, The Academic Centre, South Tees Hospitals NHS Trust

Ms Dea Birkett, Service User Representative, London

Dr Paul Cockcroft, Consultant Medical Microbiologist and Clinical Director of Pathology, Department of Clinical Microbiology, St Mary's Hospital, Portsmouth

Professor Adrian K Dixon, Professor of Radiology, University Department of Radiology, University of Cambridge Clinical School

Dr David Elliman, Consultant in Community Child Health, Islington PCT & Great Ormond Street Hospital, London

Professor Glyn Elwyn, Research Chair, Centre for Health Sciences Research, Cardiff University, Department of General Practice, Cardiff

Professor Paul Glasziou, Director, Centre for Evidence-Based Practice, University of Oxford

Dr Jennifer J Kurinczuk, Consultant Clinical Epidemiologist, National Perinatal Epidemiology Unit, Oxford

Dr Susanne M Ludgate, Clinical Director, Medicines & Healthcare Products Regulatory Agency, London

Mr Stephen Pilling, Director, Centre for Outcomes, Research & Effectiveness, Joint Director, National Collaborating Centre for Mental Health, University College London

Mrs Una Rennard, Service User Representative, Oxford

Dr Phil Shackley, Senior Lecturer in Health Economics, Academic Vascular Unit, University of Sheffield

Dr Margaret Somerville, Director of Public Health Learning, Peninsula Medical School, University of Plymouth

Dr Graham Taylor, Scientific Director & Senior Lecturer, Regional DNA Laboratory, The Leeds Teaching Hospitals

Professor Lindsay Wilson Turnbull, Scientific Director, Centre for MR Investigations & YCR Professor of Radiology, University of Hull

Professor Martin J Whittle, Clinical Co-director, National Co-ordinating Centre for Women's and Childhealth

Dr Dennis Wright, Consultant Biochemist & Clinical Director, The North West London Hospitals NHS Trust, Middlesex

## Pharmaceuticals Panel

### Members

#### Chair,

**Professor Robin Ferner**, Consultant Physician and Director, West Midlands Centre for Adverse Drug Reactions, City Hospital NHS Trust, Birmingham

Ms Anne Baileiff, Consultant Nurse in First Contact Care, Southampton City Primary Care Trust, University of Southampton

Professor Imti Choonara, Professor in Child Health, Academic Division of Child Health, University of Nottingham

Professor John Geddes, Professor of Epidemiological Psychiatry, University of Oxford

Mrs Barbara Greggains, Non-Executive Director, Greggains Management Ltd

Dr Bill Gutteridge, Medical Adviser, National Specialist Commissioning Advisory Group (NSCAG), London

Mrs Sharon Hart, Consultant Pharmaceutical Adviser, Reading

Dr Jonathan Karnon, Senior Research Fellow, Health Economics and Decision Science, University of Sheffield

Dr Yoon Loke, Senior Lecturer in Clinical Pharmacology, University of East Anglia

Ms Barbara Meredith, Lay Member, Epsom

Dr Andrew Prentice, Senior Lecturer and Consultant Obstetrician & Gynaecologist, Department of Obstetrics & Gynaecology, University of Cambridge

Dr Frances Rotblat, CPMP Delegate, Medicines & Healthcare Products Regulatory Agency, London

Dr Martin Shelly, General Practitioner, Leeds

Mrs Katrina Simister, Assistant Director New Medicines, National Prescribing Centre, Liverpool

Dr Richard Tiner, Medical Director, Medical Department, Association of the British Pharmaceutical Industry, London

## Therapeutic Procedures Panel

### Members

<p><b>Chair,</b> <b>Professor Bruce Campbell,</b> Consultant Vascular and General Surgeon, Department of Surgery, Royal Devon &amp; Exeter Hospital</p>	<p>Professor Matthew Cooke, Professor of Emergency Medicine, Warwick Emergency Care and Rehabilitation, University of Warwick</p> <p>Mr Mark Emberton, Senior Lecturer in Oncological Urology, Institute of Urology, University College Hospital</p> <p>Professor Paul Gregg, Professor of Orthopaedic Surgical Science, Department of General Practice and Primary Care, South Tees Hospital NHS Trust, Middlesbrough</p> <p>Ms Maryann L Hardy, Lecturer, Division of Radiography, University of Bradford</p>	<p>Dr Simon de Lusignan, Senior Lecturer, Primary Care Informatics, Department of Community Health Sciences, St George's Hospital Medical School, London</p> <p>Dr Peter Martin, Consultant Neurologist, Addenbrooke's Hospital, Cambridge</p> <p>Professor Neil McIntosh, Edward Clark Professor of Child Life &amp; Health, Department of Child Life &amp; Health, University of Edinburgh</p> <p>Professor Jim Neilson, Professor of Obstetrics and Gynaecology, Department of Obstetrics and Gynaecology, University of Liverpool</p>	<p>Dr John C Pounsford, Consultant Physician, Directorate of Medical Services, North Bristol NHS Trust</p> <p>Dr Karen Roberts, Nurse Consultant, Queen Elizabeth Hospital, Gateshead</p> <p>Dr Vimal Sharma, Consultant Psychiatrist/Hon. Senior Lecturer, Mental Health Resource Centre, Cheshire and Wirral Partnership NHS Trust, Wallasey</p> <p>Professor Scott Weich, Professor of Psychiatry, Division of Health in the Community, University of Warwick</p>
<p>Dr Mahmood Adil, Deputy Regional Director of Public Health, Department of Health, Manchester</p> <p>Dr Aileen Clarke, Consultant in Public Health, Public Health Resource Unit, Oxford</p>			

## Disease Prevention Panel

### Members

<p><b>Chair,</b> <b>Dr Edmund Jessop,</b> Medical Adviser, National Specialist Commissioning Advisory Group (NSCAG), London</p> <p>Mrs Sheila Clark, Chief Executive, St James's Hospital, Portsmouth</p> <p>Mr Richard Copeland, Lead Pharmacist: Clinical Economy/Interface, Wansbeck General Hospital, Northumberland</p>	<p>Dr Elizabeth Fellow-Smith, Medical Director, West London Mental Health Trust, Middlesex</p> <p>Mr Ian Flack, Director PPI Forum Support, Council of Ethnic Minority Voluntary Sector Organisations, Stratford</p> <p>Dr John Jackson, General Practitioner, Newcastle upon Tyne</p> <p>Mrs Veronica James, Chief Officer, Horsham District Age Concern, Horsham</p> <p>Professor Mike Kelly, Director, Centre for Public Health Excellence, National Institute for Health and Clinical Excellence, London</p>	<p>Professor Yi Mien Koh, Director of Public Health and Medical Director, London NHS (North West London Strategic Health Authority), London</p> <p>Ms Jeanett Martin, Director of Clinical Leadership &amp; Quality, Lewisham PCT, London</p> <p>Dr Chris McCall, General Practitioner, Dorset</p> <p>Dr David Pencheon, Director, Eastern Region Public Health Observatory, Cambridge</p> <p>Dr Ken Stein, Senior Clinical Lecturer in Public Health, Director, Peninsula Technology Assessment Group, University of Exeter, Exeter</p>	<p>Dr Carol Tannahill, Director, Glasgow Centre for Population Health, Glasgow</p> <p>Professor Margaret Thorogood, Professor of Epidemiology, University of Warwick, Coventry</p> <p>Dr Ewan Wilkinson, Consultant in Public Health, Royal Liverpool University Hospital, Liverpool</p>
--	--	--	--

## Expert Advisory Network

### Members

Professor Douglas Altman,  
Professor of Statistics in  
Medicine, Centre for Statistics  
in Medicine, University of  
Oxford

Professor John Bond,  
Director, Centre for Health  
Services Research, University of  
Newcastle upon Tyne, School of  
Population & Health Sciences,  
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 & Consultant  
Pharmacist, Rossendale

Professor Collette Clifford,  
Professor of Nursing & Head of  
Research, School of Health  
Sciences, University of  
Birmingham, Edgbaston,  
Birmingham

Professor Barry Cookson,  
Director, Laboratory of  
Healthcare Associated Infection,  
Health Protection Agency,  
London

Dr Carl Counsell, Clinical  
Senior Lecturer in Neurology,  
Department of Medicine &  
Therapeutics, 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, London

Dr Keith Dodd, Consultant  
Paediatrician, Derby

Mr John Dunning,  
Consultant Cardiothoracic  
Surgeon, Cardiothoracic  
Surgical Unit, Papworth  
Hospital NHS Trust, Cambridge

Mr Jonathan 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,  
Professor of Community  
Rehabilitation, Institute of  
General Practice and Primary  
Care, University of Sheffield

Professor Gene Feder, Professor  
of Primary Care Research &  
Development, Centre for Health  
Sciences, Barts & The London  
Queen Mary's School of  
Medicine & Dentistry, London

Mr Leonard R Fenwick,  
Chief Executive, Newcastle  
upon Tyne Hospitals NHS Trust

Mrs Gillian Fletcher,  
Antenatal Teacher & Tutor and  
President, National Childbirth  
Trust, Henfield

Professor Jayne Franklyn,  
Professor of Medicine,  
Department of Medicine,  
University of Birmingham,  
Queen Elizabeth Hospital,  
Edgbaston, Birmingham

Dr Neville Goodman,  
Consultant Anaesthetist,  
Southmead Hospital, Bristol

Professor Robert E Hawkins,  
CRC Professor and Director of  
Medical Oncology, Christie CRC  
Research Centre, Christie  
Hospital NHS Trust, Manchester

Professor Allen Hutchinson,  
Director of Public Health &  
Deputy Dean of SchARR,  
Department of Public Health,  
University of Sheffield

Professor Peter Jones, Professor  
of Psychiatry, University of  
Cambridge, Cambridge

Professor Stan Kaye, Cancer  
Research UK Professor of  
Medical Oncology, Section of  
Medicine, Royal Marsden  
Hospital & Institute of Cancer  
Research, Surrey

Dr Duncan Keeley,  
General Practitioner (Dr Burch  
& Ptnrs), The Health Centre,  
Thame

Dr Donna Lamping,  
Research Degrees Programme  
Director & Reader in Psychology,  
Health Services Research Unit,  
London School of Hygiene and  
Tropical Medicine, London

Mr George Levvy,  
Chief Executive, Motor  
Neurone Disease Association,  
Northampton

Professor James Lindesay,  
Professor of Psychiatry for the  
Elderly, University of Leicester,  
Leicester General Hospital

Professor Julian Little,  
Professor of Human Genome  
Epidemiology, Department of  
Epidemiology & Community  
Medicine, University of Ottawa

Professor Rajan Madhok,  
Consultant in Public Health,  
South Manchester Primary  
Care Trust, Manchester

Professor Alexander Markham,  
Director, Molecular Medicine  
Unit, St James's University  
Hospital, Leeds

Professor Alistaire McGuire,  
Professor of Health Economics,  
London School of Economics

Dr Peter Moore,  
Freelance Science Writer, Ashtead

Dr Andrew Mortimore, Public  
Health Director, Southampton  
City Primary Care Trust,  
Southampton

Dr Sue Moss, Associate Director,  
Cancer Screening Evaluation  
Unit, Institute of Cancer  
Research, Sutton

Mrs Julietta Patnick,  
Director, NHS Cancer Screening  
Programmes, Sheffield

Professor Robert Peveler,  
Professor of Liaison Psychiatry,  
Royal South Hants Hospital,  
Southampton

Professor Chris Price,  
Visiting Professor in Clinical  
Biochemistry, University of  
Oxford

Professor William Rosenberg,  
Professor of Hepatology and  
Consultant Physician, University  
of Southampton, Southampton

Professor Peter Sandercock,  
Professor of Medical Neurology,  
Department of Clinical  
Neurosciences, University of  
Edinburgh

Dr Susan Schonfield, Consultant  
in Public Health, Hillingdon  
PCT, Middlesex

Dr Eamonn Sheridan,  
Consultant in Clinical Genetics,  
Genetics Department,  
St James's University Hospital,  
Leeds

Professor Sarah Stewart-Brown,  
Professor of Public Health,  
University of Warwick,  
Division of Health in the  
Community Warwick Medical  
School, LWMS, Coventry

Professor Ala Szczepura,  
Professor of Health Service  
Research, Centre for Health  
Services Studies, University of  
Warwick

Dr Ross Taylor,  
Senior Lecturer, Department of  
General Practice and Primary  
Care, University of Aberdeen

Mrs Joan Webster,  
Consumer member, HTA –  
Expert Advisory Network



### **Feedback**

The HTA Programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (<http://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.***