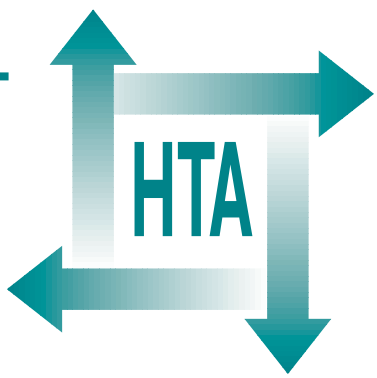
Health Technology Assessment 2001; Vol. 5: No. 8

Methodology

Issues in methodological research: perspectives from researchers and commissioners

RJ Lilford
A Richardson
A Stevens
R Fitzpatrick
S Edwards
F Rock
JL Hutton

Health Technology Assessment
NHS R&D HTA Programme



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
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). 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.

Issues in methodological research: perspectives from researchers and commissioners

RJ Lilford[1]
A Richardson[2]
A Stevens[1*]
R Fitzpatrick[3]
S Edwards[4]
F Rock[5]
JL Hutton[6]

[1] Department of Public Health and Epidemiology, University of Birmingham, UK
[2] Research Consultant, London, UK
[3] Institute of Health Sciences, University of Oxford, UK
[4] Centre for Ethics in Medicine, University of Bristol, UK
[5] English Department, University of Birmingham, UK
[6] Department of Statistics, University of Warwick, UK

[*] Corresponding author

Competing interests: none declared

Published May 2001

This report should be referenced as follows:

Lilford RJ, Richardson A, Stevens A, Fitzpatrick R, Edwards S, Rock F, *et al.* Issues in methodological research: perspectives from researchers and commissioners. *Health Technol Assess* 2001;**5**(8).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE.
Copies of the Executive Summaries are available from the NCCHTA website (see opposite).

NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS.

Initially, six HTA panels (pharmaceuticals, acute sector, primary and community care, diagnostics and imaging, population screening, methodology) helped to set the research priorities for the HTA Programme. However, during the past few years there have been a number of changes in and around NHS R&D, such as the establishment of the National Institute for Clinical Excellence (NICE) and the creation of three new research programmes: Service Delivery and Organisation (SDO); New and Emerging Applications of Technology (NEAT); and the Methodology Programme.

Although the National Coordinating Centre for Health Technology Assessment (NCCHTA) commissions research on behalf of the Methodology Programme, it is the Methodology Group that now considers and advises the Methodology Programme Director on the best research projects to pursue.

The research reported in this monograph was funded as project number 96/71/01.

The views expressed in this publication are those of the authors and not necessarily those of the Methodology Programme, HTA Programme or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work commissioned 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 editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

M

Contents

Executive summary

Objectives

- Methodological research has few well-defined tools and processes analogous to those available for reviews and data collection in substantive health technology assessment.
- This project was set up to obtain researchers' and others' views on the innovative projects on research methodology under the NHS Health Technology Assessment Programme and the usefulness of the research.
- The study was intended to span both epistemological and management issues.
- The following issues were explored:
  - the degree to which researchers would feel constrained by the "Cochrane" approach to systematic reviews when undertaking reviews of a methodological nature
  - whether methodological projects may require exceptional design and management arrangements, in view of their novelty, subjectivity and complexity
  - whether researchers would seek out other methods, in addition to undertaking reviews of argument, as a means of extending their understanding of methodological issues (there may be three categories of research methods in methodology: reviews of methodological argument, studies that use the literature as a source of data, and research that collects new primary data)
  - whether the Methodology Programme overall can be considered a "success".

Methods

- Telephone interviews were carried out on researchers (one senior and one junior per project), resulting in 35 interviews from 19 of the 20 target projects.
- A qualitative postal survey was sent to 12 people who had played a key role in the development of the Methodology Programme; replies were received from six of them.
- Analysis was undertaken of the hit rates for 29 projects on the NCCHTA website by the end of February and the end of May 1999, comparing those concerned with methodology ($n = 10$) and those concerned with other issues ($n = 19$).

Results

Undertaking methodological research: views of researchers

This section summarises the views of 35 researchers who were interviewed by telephone.

The nature of methodological reviews

- There was a reluctance among researchers to use the term "systematic review" in the methodological context.
- Practical problems in undertaking methodological reviews were found at every stage of the research process.
  - In the initial search stage, preplanned strategies were difficult to maintain, owing to the need to respond to the problems of too few or too many references.
  - At the analysis stage, most studies were not formally weighted, but there was implicit weighting in researchers' views of their merits or relevance.
  - It was often only at the synthesis stage that researchers could see clearly what their study was able to do; iteration was frequently necessary at this point.
  - It was difficult to form simple conclusions and recommendations beyond summaries of what was known in the field.
  - Dissemination activities were most often directed to other health service researchers, with some attention to NHS policy makers and research commissioners.

The need for flexibility

- Few researchers had amended their topic or methods once their research was under way, although some had made minor changes to their original plan, generally to refine the topic to fit the time or data available.
- Changing a topic was seen as inappropriate unless checked with funders, but changes in research methods were viewed as reasonable because questions might be refined in the light of information gained or early thinking.

The question of bias

- Few researchers considered that this kind of research could be undertaken or presented

in a wholly unbiased way because of the need to assess the research studied.
- Objectivity was nonetheless seen as something that researchers should strive towards. Efforts to do so included presenting data clearly, separating findings from discussion, covering all points of view, setting out their own assumptions and values, and testing their ideas on others known to have differing views.
- The formal peer-review process was not seen to have made a difference here, primarily because of the stage at which referees become involved.

Project management

Timing and time management
- A majority of projects were completed within 3 months of their due date. Those studies completed roughly on time were considered to have efficient junior researchers and good project management, including clear deadlines for different stages of the research.
- Some studies had severe problems of time management. Too much time tended to be spent on collecting and reading the literature and the writing stage was not always well planned. Referees' comments were also slow in coming.

Day-to-day management
- Most projects were considered to have been well managed, but some had serious problems.
- Good management consisted of a clarity of roles between senior and junior researchers, active supervision of the latter and a set timetable. Careful time management was seen as essential, including building in a sufficient amount of a senior person's time.
- Problems noted in the projects included staffing difficulties (both unanticipated departures and researchers' unsuitability for the work), being under-resourced and the project not being well planned from the outset.

Involvement of others
- Grantholders were not always actively involved, but this was not generally seen as problematic. There was no consensus on whether payments to senior researchers should be built into projects.
- Only a few projects had a steering or advisory group beyond the grantholders. Such groups were generally viewed as helpful because they served as a sounding board and brought in additional expertise.
- Most researchers thought that the HTA Programme itself had exercised an appropriate level of involvement.

Collaboration
- Many projects overlapped with other HTA funded projects and entailed some collaboration among researchers, which was generally seen as helpful in saving time and in stimulating ideas.
- Some questioned the desirability of collaboration across similar projects because separate projects could develop independent corroboration of results; there were also problems of academic rivalry.
- There was much less collaboration with groups outside the HTA Programme.
- The Methodology Projects Group, an informal meeting of grantholders and research fellows, was widely seen as helpful as a means of learning about other projects and obtaining moral support in the face of difficulties.

Referees
- Most researchers considered that the referees' comments had been helpful by providing intellectual support, forcing them to rethink their arguments, or pointing out gaps in the research.
- Referees were generally believed to have had the right skills, comprising a mix of specific professional and methodological expertise.
- Most respondents thought that projects needed two or three referees, although those with more complex projects tended to consider that more referees were needed.

Issues for the HTA Programme
- Some considered that methodological questions could be bolted on to other research, such as clinical trials.
- The view was expressed that methodological reviews should be updated as needed. Those who carried out the initial review should be asked first because they could do such work efficiently.
- Opinion was divided on whether the Programme should fund more large studies or small projects designed to scope a topic rather than deal with it comprehensively. There was also no marked preference for more primary or secondary studies, as this depends on the questions asked.
- A number of ideas for future research were proffered, but there was also interest in more attention being paid to getting the results of research into practice.

The Methodology Programme: views of those involved in its creation and development

This section summarises the views of key people who responded to a letter containing three principal questions.

Aims of the Programme and their achievement

- The aims of the Programme were seen to be identifying and answering important methodological questions relevant to health technology assessment and raising the profile of health services research methodology.
- Four respondents thought that the Programme had been highly successful, with impressive output in terms of both quantity and quality. This was ascribed to enthusiasm from researchers, assisted by the Methodology Projects Group, and good steering from the Methodology Panel itself.
- Doubts about the Programme included whether the right research had been funded and whether the research had been well managed by the commissioning body and research teams. Two respondents considered that a shortage of trained researchers had hampered the development of the Programme.

Noteworthy projects

- Two interviewees identified the characteristics of a successful project. Both focused on the qualities of the researchers: there is a need for the involvement of senior researchers who are familiar with the field and the Programme, and who are able to think deeply about the research and understand the target audience. A multidisciplinary team was also seen as important.
- Various projects were noted as having been well executed, often following the particular interests of the respondent. Two noted that the projects concerned with randomised trials were particularly valuable.

Future directions for the Methodology Programme

- It was argued that the dissemination and use of methodological research needs to be addressed.
- Attention to research management was seen to be needed. Projects should be required to involve senior staff and the submission of early drafts of the final report for discussion.
- There was concern that the decision to widen the portfolio beyond health technology assessment would bring new problems of defining both the content of the work and who the customers are; in consequence, an evaluation of the Panel when its remit widens was proposed.
- Other suggestions included the need for systems for updating reviews and to address the problem of systematicity in the context of methodological work. It was proposed that methodological gaps in the Service Delivery and Organisation agenda should be identified and new primary research commissioned.

Interest in the projects

- On the basis of "hit rates" for the relevant website, there is considerable interest in these projects. The monthly hit rate for methodology projects was not only high in itself (median 264/month) but was also more than twice that of other HTA projects (median 102/month).
- More recently published projects were found to have higher hit rates.

Conclusions

Reflections on the findings

- These studies were commissioned at a time of high enthusiasm for systematic reviews and meta-analysis, resulting in both epistemological and practical problems for some of those undertaking them.
- The importance of systematicity rather than exhaustiveness needs to be recognised.
- The question of whether these reviews should be seen primarily as data collection or thinking exercises pervaded the interviews and has important practical implications.
- Contrary to our expectations, little was made (in this sample) of additional research methods, with few attempts at triangulation. Researchers were not opposed to changing methods in principle but did not do so in practice.
- Researchers seemed very conscious of the problem of bias and undertook an impressive range of steps to reduce its impact.
- The organisation and management demands on these projects were not notably different from those for other studies, including the need to plan carefully from the outset and for close research management by senior staff, but some issues may be heightened in this context.
- The Programme was largely seen as a success, covering a wide range of issues and helping to develop a pool of researchers familiar with the field.
- The website analysis provided a quick illustration of the considerable interest shown in methodological and other reports.
- The significance of the Methodology Programme should be seen to lie not simply in the reports produced, but in the diffusion of knowledge it facilitated.

Reflections on the methods

- The telephone interviews worked well, although they were time-consuming and tiring; they elicited very full responses and much valuable material.

- The letters to key people were clearly less successful, with a very poor response rate. Those who did reply generally provided very thoughtful responses.
- The website analysis was a useful addition to the research but, like citation analysis, cannot be used to infer the quality or impact of the research.

Recommendations

Recommendations for researchers

The conduct of research

- Researchers should consider methods beyond the review of ideas and even the review of data, for instance, networking and other means of primary data collection (e.g. methodological studies attached to primary health technology assessment).
- Because systematic reviews in this context are very different from traditional "Cochrane type" reviews, methodological researchers should not try to chase every reference, but ensure that they search widely (i.e. consider disparate databases and sources).
- Some overlap of the various stages of research – searching, analysis, synthesis and writing – should be encouraged because this can help to clarify the nature of the research.
- Researchers should publicise their studies early on, to help to short-circuit extensive search processes and stimulate ideas.
- All studies should include a short summary of key findings, which should include practical solutions to identified problems, to assist future researchers.

Reducing bias

- A variety of safeguards to reduce potential bias should be built in to research, including the establishment of a steering group, multidisciplinary teams, peer reviewing of applications and final reports, a report for the commissioning body, and a clear intention to publish in widely disseminated journals.

The efficient management of research

- Senior staff need to be closely involved throughout the research, both to assist with data analysis and to maintain good research management; this should be reflected in the costing of grants.
- A clear plan of action and research timetable should be developed, including a plan for the report writing, with a preliminary structure to influence both data collection and analysis.
- Multisite projects should be considered carefully before they are set up, with respect to both the logistics and the willingness of the parties to work together.

Recommendations for the Methodology Programme

Project management

- The Programme should continue the fairly light style of research management that it has used to date, but remain accessible to researchers who may need to discuss problems or changes to the initial plan. Programme representation on a project steering group is one means of achieving this.
- The Programme also needs to continue to advise researchers on overlapping projects; care needs to be given to avoiding duplication of effort through discussions with project advisory groups. A master steering group for all projects would be one way of achieving this and reassessing resource needs.
- Prior to commissioning projects, the Programme could introduce a more iterative approach, to ensure that it has correctly specified the research problem.
- Particular attention should be given to the dissemination and use of the findings of research already commissioned by the Programme and others; websites are clearly valuable for this purpose.
- The need for mutual support among researchers undertaking complex projects should not be overlooked.
- The new Programme could be evaluated to see if it continues to work as well as it did when its brief was more limited.

Future commissioning

- Methodological researchers should be encouraged to explore a wider range of methods, incorporating intellectual analysis and primary research, including methodological experiments.
- Some substantive researchers should be encouraged to add a methodological component to their studies, but not all subject areas can work within such a framework.
- The management of methodological research may itself be the subject of study.
- Arrangements could be set up for updating reviews as needed.
- Particular attention could be given to methodological gaps in the Service Delivery and Organisation agenda.

Chapter 1
Introduction

In 1991 a major step was taken in establishing a research and development (R&D) strategy for the NHS. There were few, if any, precedents for a health service seeking to identify its research needs in a systematic and concerted way, rather than relying on scientific investigators' priorities. The objective of the strategy was to provide a knowledge or evidence base, not just for clinical but also for policy and managerial decisions within the NHS. Subsequent initiatives in the NHS have underlined the importance of the R&D Programme that emerged. The recently established National Institute for Clinical Excellence, for instance, requires evidence about clinical effectiveness and cost-effectiveness because one of its principal functions is to improve the quality of NHS services by issuing and disseminating evidence-based guidance. Similarly, evidence from the R&D Programme has played an important part in the introduction of national service frameworks, with their concern to overcome geographical variations in NHS services by setting national standards for particular patient groups or services.

The majority of work commissioned by the NHS R&D Programme has been substantive, that is, research addressing specific clinical or policy questions. Another unusual – if not unique – feature has been to establish a programme to develop research methods for evaluating the effectiveness of healthcare interventions. Thus, a Methodology Panel was set up within the Health Technology Assessment (HTA) Programme of the NHS R&D Programme, whose purpose was to prioritise work to develop research methodologies for the HTA Programme as a whole. This has since been turned into a programme in its own right. For ease of discussion, this is referred to as the "Programme" throughout this report.

Altogether, roughly 40 projects have been funded by the Methodology Programme to date. The overall aim of this activity has been to develop a corpus of methodological knowledge and analysis to underpin and strengthen the HTA Programme as a whole. Much of this work has involved "systematic reviews", covering such topics as Bayesian statistical methods in clinical trials, assessing fast changing technologies, and the ethical aspects of clinical trials. Most projects were set up to run for 1 or 2 years, costing about £50,000.

Methodological debate

The role of the Methodology Programme within NHS R&D may easily be underestimated because the greatest controversies and more visible impact of research tend to arise naturally in relation to evidence of the effectiveness (or lack of effectiveness) of specific interventions. Evidence from evaluations of specific interventions is primarily intended to inform judgements about their value and, ultimately, the appropriate extent of public provision. Considerable professional, scientific and, indeed, media and public attention is therefore devoted to the substantive results of evaluative studies, especially when such evidence strongly indicates the value (or lack of value) of a drug, form of surgery, diagnostic technique or other form of service.

Methodological research, by contrast, may seem more removed, abstract and less policy-relevant to "real world" decision-making in healthcare systems. To some, the role of a methodology programme may therefore be considered to be to address secondary, technical and specialised questions intended to influence the quality of subsequent substantive research, while having little direct or immediate importance in its own right.

There are, however, several reasons for arguing that methodology may have greater significance – and therefore warrant more direct attention – than being solely a technical resource for more interesting "real" research. First, the research methods at the heart of current calls to increase the NHS evidence base are not without controversy. As a number of authors have argued,[1,2] at the risk of over-simplification, "evidence-based medicine" is shorthand for a distinctive emphasis on statistical, epidemiological and probabilistic models of health and healthcare, which contrasts with and, perhaps, challenges conventional biomedical methods. Indeed, health professionals do not readily or automatically accept and act upon evidence-based medicine, in contrast to other sources of information about practice.[3]

A second reason for paying greater attention to methodology as a subject in its own right is that apparently technical debates within healthcare

evaluative research extend beyond the purely technical. For example, the merits and role of non-randomised evidence for healthcare interventions are considered very modest within the methodological paradigm of evidence-based medicine, which gives greatest weight to well-conducted randomised trials and meta-analyses of such trials. In epidemiology, other study designs are used. Debates about the merits of observational evidence, however, although expressed in technical and methodological terms, can often appear to reflect more basic conflicts. In particular, those who suggest that evidence-based medicine be entirely limited to randomised controlled trials can be considered to be taking an absolutist view. In short, the position one takes on the value of data from different health service research designs will affect the final interpretation of the technology in question.

Lastly, and probably most importantly, the methods used in evaluative research increasingly stray into less purely technical and potentially more contested domains. The NHS R&D Programme extended the debate about health care interventions beyond the narrow confines of clinical effectiveness to address the costs and broader impact of health technologies. These issues must be assessed in order to inform NHS decision-making. Even if questions about the best methods of evaluating clinical effectiveness were resolved, purely technical solutions are not close to being reached on the methods whereby the costs, value and social impact of health technologies can be authoritatively assessed.

The NHS R&D Programme extended into territories where issues of human value and choice must be considered. It looked to the Methodology Programme to bring about improved methodological quality in the research that assesses costs, values, preferences and social impact. Whether or not it is, in principle, feasible to address issues of social value with the same precision that appears to be expected in evidence of clinical effectiveness, it is at least likely that the methods involved in such research will be subject to greater debate.

Methods in methodology

The project reported here was designed to obtain a better understanding of the research process undertaken by contributors to the Methodology Programme. There were several reasons to examine the development and early evolution of that Programme. First, as the first Director of the HTA Programme pointed out, investment in the Methodology Programme was identified early on as a priority for NHS R&D.[4]

Secondly, despite a vast array of methodological literature and research practice on which researchers could draw, it was not obvious how a programme of methodology research should be conducted, given its innovative nature in international terms. A core of methodological assumptions, which can be summarised here as evidence-based medicine, probably informed early expectations of the Programme, especially regarding the merits of randomised evidence and particular methods of collating and synthesising evidence. It was not easy to estimate how well the paradigm of evidence-based medicine would be sustained if rigorous and questioning methods were used to examine and test such assumptions; nor was it easy to judge how well the paradigm would address questions beyond the field of clinical effectiveness in which it had been developed.

At the inception of the Methodology Programme, it was taken for granted that the methods of evidence-based medicine would prove both useful and necessary in gathering and assessing evidence for these methodological systematic reviews. It was not questioned whether the methods of methodological research were in any way problematic or insoluble, given the positive perceived impact of Cochrane-standard reviews that underpinned evidence-based medicine. Thus, it is of considerable interest to examine the substantial and innovative investment made by the NHS in developing the methods whereby health services are evaluated.

This study was based primarily on in-depth interviews with researchers who had undertaken a study of a methodological issue. In addition, the views of a few key people involved in commissioning such reviews were sought and the degree of interest in such reviews, as evidenced by the use of a website, was examined. Its overall aim was to understand how such research was and could be undertaken, and what could and should be expected from it, as well as how such research could be made most useful in future. Attention was given both to the process of undertaking such reviews (e.g. the nature of hypotheses, appropriate search strategies) and to the organisation and management of the projects as a whole. No attempt was made, however, to assess the quality of the reviews studied, as this was seen as beyond the brief of the research.

Hypotheses

The members of the research team had all undertaken methodological research for the Programme

and had an interest in exploring how such research may be undertaken, so we had certain hypotheses and expectations for our findings. First, it was our expectation that researchers would not feel bound to undertake methodological reviews by Cochrane Collaboration methods, although they could be considered the default. Our view was that methodological reviews were somewhat different from standard systematic reviews, generally searching for ideas rather than new data, although not regressing to undisciplined narrative. The probability of uncovering new ideas in a review reduces as the search progresses. The methodological literature is very extensive and quality standards to define plausible cut-offs are not well defined, so it seemed unlikely that most researchers would seek to be exhaustive. It can be added, indeed, that by the time the first tranche of methodological reports (not all reviews) were completed, the HTA Programme itself was no longer using Cochrane Collaboration terminology to commission new methodological work. It now distinguishes methodological reviews from systematic reviews.

Secondly, the research team had other expectations for the organisation and management of methodological research, which guided the design of the questions asked. These concerned the particular needs of methodological research for iterative discussion, seniority of researchers, multidisciplinary working, the use of outside experts, the difficulties of being value free, and uncertainties about the needs of the audience.

Thirdly, the methods for methodological research were not well defined. Given the intellectual interest of researchers in their own methods, we considered it likely that they would wish to explore possibilities beyond reviews (Cochrane or otherwise) to feed their research. At its simplest level, this could mean adding to written literature with word-of-mouth surveys of innovative researchers in the field and intensive discussion with accessible key players. In a sense, such methods are equivalent to the gathering of new data in substantive (non-methodological) research.

We considered that there may be three clear categories of research methods:

1. Strict reviews that summarise and analyse existing methodological argument. In effect, these are the methodological equivalent of Cochrane Collaboration systematic reviews, although the constraint on exhaustiveness would make them distinct.
2. Methodological research that uses the literature as a source of data (rather than reviews it). Such research, for example, may be concerned with the potentially different outcomes of studies using various methods to approach the same substantive HTA/evidence-based medicine problem. An example would be studies that compare the outcome and apparent value of a cohort study and a randomised trial of the same technology.
3. Research that collects new primary data, whether by survey or, exceptionally, the undertaking of new substantive primary research by more than one method to evaluate the relative merits of different approaches.

We recognised the potential value of discussions within a methodological team in generating new ideas, but this can be considered to be a special case of the survey of experts method (point 3 above).

The team expected that particular studies would not be limited to any one of the approaches above, but could well use more than one and, indeed, triangulate to their conclusions by the use of more than one method. Our expectations that the classification would be vindicated by our research were mitigated by the knowledge that our sample was limited to the first year of the Methodology Programme, a year in which the review paradigm was all but assumed. This was partly owing to the ascendance of the Cochrane Review methodology and partly to the fact that those initiating the Programme, quite rightly, favoured bringing together what was already known in the literature, prior to funding new primary research. Subsequent years of the HTA Methodology Programme have widened this framework.

Finally, we had an open mind about the usefulness or otherwise of the research findings. We chose to explore these with key players and potential users of the Programme's output.

A full list of all the projects studied is provided in appendix 1.

Chapter 2

The research methods

The methods used for this study comprised a telephone interview survey of researchers, a small postal survey of key people involved in commissioning the Programme, and an examination of the interest in the results as evidenced by use of the HTA website.

All grantholders were closely involved in the development of this work via meetings of a small working group. We consider that it is important to acknowledge the significance of this group because the discussions were an important stimulus to thinking deeply on issues raised by this research.

Survey of researchers

The interview schedule was developed from issues raised in the protocol, revised and refined in response to grantholders' suggestions. It was piloted initially with two researchers undertaking reviews (including one from those commissioned by the HTA in its second year (i.e. the year after the projects in this study)) and, after revisions, with two further researchers from the study period. The interviews were undertaken by the research assistant over a 2-month period (mid-July to mid-September 1998). A copy of the interview schedule is provided in appendix 2.

It was our original plan to interview both a principal investigator and a research associate from each HTA methodology review commissioned in 1993 (20 projects in total). We initially contacted both individuals for each study, to introduce the research and alert them to the project's interest in an interview. The principal investigators were subsequently sent details of the purpose of the project and the interview, including subject headings and its likely duration. Interview dates were arranged by telephone.

The interview itself, conducted by telephone, began with a check that the interviewee was aware of the expected length of the interview and his or her availability to devote sufficient time. Consent to recording was also obtained. The questions were then asked, generally in the order of the initial schedule. On occasion, one was omitted when the interviewee had answered the question very exactly already but, if the issue had been discussed only briefly, the question was repeated when it came up in the interview schedule.

In all, 35 interviews were undertaken, spanning 19 projects. Of these, 18 were with principal investigators and 17 were with research associates; for one project it had proved impossible to obtain an interview with anyone involved and, in three others, only one researcher could be interviewed. The interviews lasted about an hour, although some were considerably longer. In one, insufficient time had been allowed by the respondent and, because of difficulty in scheduling a follow-up interview, the questions were sent by e-mail and responses received by this medium. With the inevitable lack of probing, these answers tended to be shorter, but we thought it was better to obtain some response to the remaining questions rather than none at all.

The tapes from the interviews were transcribed verbatim and given on disk to the researcher. Analysis was undertaken by first creating new computer documents, comprising each researcher's response to sets of related questions, laid out in project order. When information was provided in the "wrong place" (in response to a different question), every effort was made to put such information with the appropriate question. The analysis could then be undertaken very readily, coding answers as appropriate to each question.

All interviews were read through carefully and the responses analysed; every attempt was made to avoid the selective use of any findings. Where appropriate, the responses were analysed by project. Where there was some potential relevance, the responses of interviewees in specific subgroups were compared, for instance, the responses of senior and junior investigators. In some cases, only certain subgroups could be analysed, for example, those who had been through the refereeing process. Wherever possible an attempt was made to quantify answers, so that a view of the relative weight of the responses could be obtained. As with much qualitative research, however, this did not work fully in practice because the respondents did not invariably answer a question directly or at all.

The following discussion sets out this information as fully as possible, but some shorthand is employed. For instance, it is simpler to state that "most" researchers took a particular view, without belabouring the fact that it was most of those who answered the question.

Survey of those with a key role in the Methodology Programme

Initially, we had intended to undertake a fairly sizeable survey of potential and actual users of methodological research, to obtain their views on the utility of the research commissioned under the Methodology Programme. A number of practical difficulties presented themselves, however, including the extent to which they would invariably assess the appropriateness of individual projects rather than the outputs of the Programme as a whole. We concluded that a more modest exercise was in order.

The second part of this study therefore comprised a letter to 12 people who had played a key role in the development of the Methodology Programme. These included: the R&D Programme's Director at the start of the Programme; the Deputy Director of NHS R&D at the time; the Chairman of the Advisory Group on health technology assessment; the two HTA Programme Directors; the two Chairmen of the Methodology Panel; the two Chairmen of the Commissioning Group/Board and the three successive secretaries to the Methodology Panel. Their names are set out in appendix 3.

These key respondents were asked: to comment on the usefulness of the Methodology Programme, giving their reasons; to note any particularly valuable projects; and to suggest how the Programme should develop in future. They were asked to base their assessment in part on the perusal of a text[5] providing a useful summary of the output of the methodology work to date. Those who did not already have a copy were sent one on request.

Regrettably, this part of the research was not as fruitful as we had hoped. Despite repeated reminders by telephone and e-mail, only six of the 12 responded. It must be noted that, because many of these people were known to the grantholders, all were promised complete anonymity regarding both the content of their replies and whether or not they replied at all. This seemed essential in order to elicit an open and honest response in such circumstances. The only person with this knowledge was the external research consultant, who did not know them personally and had no reason to be working with them in future. She was careful not to divulge it subsequently. The letter used is set out in appendix 4.

Analysis of the HTA website

In order to obtain some, albeit highly imperfect, measure of interest in the results of the 1993 HTA Programme, we sought information from the NCCHTA website on the number of "hits" scored by each published project. We used the recorded data on the number of hits for 1993 projects at both the end of February and the end of May 1999, as well as noting the month of publication of each project. Projects were then classified according to whether they were commissioned under "methodology" or "other" (including acute sector, diagnostics and imaging, pharmaceutical, population screening, and primary and community care).

In all, we found ten methodology projects posted on the website and 19 non-methodology reports. We were aware that the data source was regrettably incomplete at the time of the exercise. The missing data (identified from the Annual Report of the NHS HTA Programme 1999) consisted of four published reports (three methodology and one other) and 41 projects (seven methodology and 34 other) that had not yet been published. Among the latter, the great majority were still under way (six methodology and 31 other), but three had been rejected (one methodology and two other) at the point of submission or had been discontinued for some reason. In addition, two non-methodology projects had been amalgamated.

Summary statistics and scatterplots were developed using these data and regression analyses were then carried out.

Chapter 3

Undertaking methodological research: views of researchers

This section summarises the views expressed in discussions with 18 senior and 17 junior methodological researchers who worked on 19 research projects, obtained through telephone interviews.

The nature of methodological reviews

A central focus of this study was to explore the practical problems of undertaking systematic reviews in the area of methodology. Without question, such problems pervaded most interviews because few researchers had found the process to be straightforward. Although some admitted to such problems with considerable reluctance, most were very ready to discuss them in great detail and explain why they arose.

The following section considers the various stages of the research process, noting where appropriate the dissimilarities from a traditional systematic review.

The "type" of study

The first way in which the reviews studied differed from systematic reviews was in the terminology used to describe them. With a few exceptions, the researchers found it difficult to define the "type" of study they had undertaken. Considerable discussion took place about definitions of "systematic review", but generally it was believed that their studies did not qualify. One commented: "It jars in my feeble brain to call it a systematic review."

The great majority (24 researchers) argued that they had begun with the intention of doing a systematic review or had attempted to be as systematic as possible, but that their research could not properly be described in this way, at least "not in the Cochrane sense". Various compromise terms were offered, such as "quasi-systematic" or "hybrid", or they suggested that their aims were to be "thorough" or "authoritative" or "explicit about how the literature was searched". Some researchers proposed alternative terms for their work, such as "comprehensive review", "methodological review", "mapping review" and "narrative review", although the latter was seen to have pejorative overtones.

One researcher, suggesting the terms "thoughtful" or "scholarly review", contrasted it with a traditional Cochrane style review, which was seen to be "not very thoughtful":

> "In fact, it's thoughtless – it just makes some bizarre assumptions, simplifies and reduces everything to counting. Where is the actual scholarship and thinking that is brought to bear?"

Only six respondents declared with certainty that they had undertaken a systematic review. These were mostly junior researchers and none was from the same project.

When pressed to describe their studies, a small majority of researchers ($n = 19$) said that they were "systematic but not exhaustive" (one added "and exhausting"). It was argued that there was no point in trying to cover the literature comprehensively because there were diminishing marginal returns in finding yet another exposition of the same argument: "After a while we were just hearing the same things over again." This was contrasted with studies involving meta-analysis, where there was a need to track down every study:

> "If you've got five trials and you find a sixth trial on something, then it's usually quite useful to know what the sixth trial's got to say. But if you've got five methods papers, finding a sixth might be useful or it might be … completely unuseful. So they are different things. This is the problem with thinking about it in terms of a systematic review."

> "With a methodological review, as long as we get a large proportion of the literature, then we should get a feel for every single method that is out there. We don't need every paper to ensure that we discover every method that is being used … If we had done a totally exhaustive search, we would probably have ended up with some extra papers. But … it wouldn't have added anything in terms of the content of the report."

For others, being exhaustive was seen as an impossibility, as discussed in more detail below. On the other hand, seven researchers believed that they had been exhaustive.

Some anger was expressed at the ways in which their task had been defined. It was said that the research commissioners did not understand the nature of the task in this context:

> "The dialogue with the NHS R&D Executive was very constrained by the desire to turn everything into a systematic review … which [for] this study was completely inappropriate … At the end of the day, it contains rather more systematic review than it ought to and rather less qualitative exploratory research than it ought to from the point of view of good methodological design."

> "Doing a systematic review of methodology is not the same as doing a systematic review of substantive studies in an area. The commissioning process … ought to recognise that. The existing mould of how you do a meta-analysis of 30 or 300 clinical trials – some aspects of that are relevant to a methodology review and some are not. And to force everything into that mould is very counterproductive."

Expectations at the outset

Differences from traditional systematic reviews can be seen right at the early stages of the research process. The great majority of researchers ($n = 24$) did not have firm expectations regarding their findings or, indeed, hypotheses to be tested. Instead, they had expectations of the process and the expected outputs. The latter were generally viewed as a report describing and analysing the relevant literature ("a mapping exercise"), often making firm recommendations for policy makers, researchers or research commissioners:

> "It wasn't answering a research question in that sense … I am involved in studies to test the hypothesis of if this treatment works or if that strategy is a good idea. [This research] was more of a descriptive study. It had objectives, but not a hypothesis."

Several respondents had also planned to produce checklists or guidelines.

A number of researchers added that they had some clear expectations of what the literature would contain or what the key issues were. More often than not such expectations were confirmed, but some surprises were noted. Several found that the relevant literature was more limited than they had anticipated and the quality, in terms of the depth with which issues had been explored, was felt by one or two to be poorer.

Search strategies

The process undertaken by these researchers also deviated from a traditional review in their search strategies. Although most projects (14/19) were described initially as having a preplanned strategy, it became clear with probing that, in the majority of these ($n = 10$), this was substantially modified over the course of the work. In addition, five projects were said by researchers to have no preplanned strategy at all.

Of the four projects in which there was no deviance from the original plan, the terms and databases were seen as very straightforward. This was helped by the employment of a qualified librarian in two cases.

Much more commonly, researchers spoke of using "trial and error" or, more formally, "an iterative approach", as they learned from what they were doing. To some extent, the approaches described were very much the same. Researchers would devise a set of key search terms and see what they retrieved. The process of devising them tended to be by brainstorming with the research team (and sometimes others). The databases also tended to be similar (primarily MEDLINE, EMBASE, BIDS and occasionally Cochrane), depending on the study. Many researchers also supplemented their search by writing to colleagues working in the field, exploding references and undertaking manual searches. One also tried using the Internet to supplement other search strategies.

The reasons for using an iterative approach were not all the same; the difficulties experienced varied across studies. In five projects, researchers found too few papers for all or some parts of their study (some projects entailed more than one part, with different search experiences). These respondents tended to liken their task to finding a "needle in a haystack" and generally found it very frustrating, although some expressed this as a challenge:

> "The people who had sent out the call for proposals didn't seem to be aware that you can only do a systematic review of … areas that have been extensively studied. If you don't have a literature to review, you can't do a systematic review satisfactorily."

> "There was one particular topic where not much had been written at all, but that made it quite interesting … Because there wasn't much literature, we had a stab at writing something ourselves on it."

These researchers had to be particularly inventive in their search strategies. They searched databases for authors who were known to write in the field, looked for keywords used in articles found, or simply contacted others known to be working in the subject area. They also tended to explode references and search the grey literature, including conference literature and textbooks, and to handsearch the key journals or other sources.

> "We just kept redefining it. We did it dozens and dozens of times; we had to keep modifying it different ways and creating a matrix of citations and different search strains of what modifications we were doing in terms of hit rate …"

Some also found that, for their subject, indexing was so poor that they kept going down blind alleys:

> "There's so much misclassification in the whole indexing system; it just turns up so much rubbish at the end of the day."

> "We ran into completely inconsistent terminology: 'longitudinal studies' have very specific statistical meaning, but can mean any kind of follow-up study, ranging from 3 days to 20 years, rats or humans. That work was very, very difficult."

More than half the projects ($n = 11$), however, generated too many articles in the first instance. There were some real tales of woe concerning the outcomes of initial searches; at the highest end, researchers spoke of retrieving 10,000, 25,000 or even 60,000 articles on an initial search. As with finding too few articles, this also required rethinking and some inventiveness to make the task more manageable. Some clearly felt "overwhelmed" by the amount of information found:

> "It was unfortunately months rather than weeks, when we decided that we had to be quite brutal about this … We did get lost a little bit in the amount of information that we were generating."

Many researchers, including some who initially said they had too many articles, thought that the number was "manageable". To some extent, such a judgement may depend on the personality and expectations of the researcher, but a key factor was whether the articles to be reviewed could be filtered at the abstract stage. A project that retrieved 2,000 articles, all of which had to be read on hard copy, had much more difficulty than one that retrieved 3,000 that could be whittled down by reading the abstracts. It also depended, of course, on the final number of articles found and the clarity of the inclusion criteria. For some, the process was relatively straightforward:

> "There was an awful lot of stuff which just rehearsed arguments that had been made before, that aren't adding any new intellectual material. And because we came to know those arguments well, it didn't take us long to scan the articles and say 'Yes, there's nothing new in this.'"

> "It's not so daunting – if you're reading a paper about an application and you're very focused on looking for the methodological bits, you don't have to read the paper from cover to cover. A lot of it's just skimming."

The difficulties of having too many articles were of a number of different kinds. For some, the key issue was the sheer time needed to read through any one article:

> "Methodological papers are just horrendous to read. If I am refereeing, I can read a medical paper and decide what it is saying … in an hour or more. If I referee a statistical paper, it will typically take me the best part of a day. So the idea that we could have read and absorbed anything like those numbers of methodology papers in retrospect was barmy."

Some developed a strategy to cut down the obligation on them:

> "There was, I suppose, an informal strategy that said if you get an additional ten papers and there are no new arguments, you start to be suspicious. And if you get an additional 100 papers and there are no new arguments, it is probably worth stopping."

For others, the breadth of the area covered made it hard to cut the task down to manageable proportions. One strategy in this case was to limit the project to a specific time frame, a specific discipline (where the project could be seen to cover more than one), or to concentrate on a subset of the original topic. Another was to introduce one or more screens to filter out completely irrelevant material, using the abstracts and sometimes duplicate reviewing of decisions for extra care.

> "There is often a need to infuse a bit of common sense into the whole process of doing literature reviews … You've got to say 'What is the question you're trying to answer?', not 'Can we find out everything we want to know about this topic?'"

> "We went through everything we retrieved and sorted out the completely irrelevant articles … and then we used another screening where we tried to look really hard at the key issues and the relevance to this project. We boiled it down to a few thousand references."

One researcher described setting up such a complex strategy that the computer kept crashing online.

Another problem was finding many articles on the same subject, but with only a small proportion of any one relevant to the purpose or containing sufficient detail:

> "We ended up with a stack of papers that … didn't go in-depth enough for what we wanted, but we couldn't get rid of them out of the search."

> "Some articles addressed the issues that we were interested in, but they were simply illustrative examples … that didn't contribute anything new. In some ways, they are the most difficult to get through … They used the same method that 20 other people have used, but some of them do make slight modifications, which you want to be aware of."

Some researchers were constrained by other factors, such as resources (for instance, when they had to

pay for interlibrary loans), access to databases or articles, or the fact that much of the relevant literature was not retrievable through electronic search methods.

Most (n = 21) admitted that they found some references by chance; only nine were certain that they had never done so. These references came via a variety of routes: books or articles that they happened to own, being asked to referee a relevant article, or leafing through journals for some other purpose, but most often from comments from colleagues (either local or in other institutions). Some researchers queried the term "chance" because they had made conscious efforts to improve such finds through informing colleagues about what they were doing.

When asked how such chance finds could be reduced, some researchers spoke of better indexing, especially of grey literature, including PhD theses and conference papers. A simpler suggestion was that all researchers should try to publicise widely their ongoing research for just this reason. It was also noted, however, that this would inevitably involve some subjectivity and that indexing methodological issues was bound to be problematic when they were not the primary interest of the initial author.

Finally, the occasional researcher found the whole process to be fun:

> "I am a bit of a hoarder/collector and to be told to go out there and collect everything that has been written of substance on a subject – and to have a set – was very appealing. When you trace these really tentative links, you feel slightly on a detective case …"

Analysis and synthesis

Of all the stages in the research process, the activities of analysis and synthesis were the most difficult for researchers to explain. Again, such activities also tended to differ from what would be undertaken for a traditional review because a much more qualitative approach was required.

First, no project (with the exception of one) formally weighted the studies collected. The researchers argued that such weighting would be inappropriate for the kind of study they were undertaking:

> "We were just interested in identifying issues, so there was no point in identifying [projects] that actually conduct a study better."

> "Weighting wasn't really relevant … We weren't doing a meta-analysis, we weren't trying to come up with a figure for a particular treatment effect."

Some also noted that weighting would be difficult because of insufficient research in the area or for other reasons:

> "At the moment there is no checklist or agreed way of deciding … and although the methods are used a lot, they're rarely written up in any rigorous way methodologically to be able to judge how well it's been done."

> "When you're doing philosophical research, it's not so simple, because the way in which an argument is put forward can be very variable – a terrible paper can have a good idea in it and so on."

On the other hand, many researchers suggested that there was an "informal" or "implicit" weighting in which studies they chose to include. First, they had to take decisions on what was deemed relevant and, secondly, they tended to take into account the merits of the research involved, sorting the "truly terrible" from the "sensible":

> "By discarding a lot of stuff, you are obviously weighting it as having zero value … It was entirely informal – judgements that we either individually or from collective meetings made about whether someone was saying something important, original, worthwhile, influential."

> "We did implicitly weight papers by selecting the papers that seemed to be most relevant and contributed the most to a particular subtopic … There are a lot of papers that didn't make it even to the annotated bibliography."

Several respondents also said that they commented in their reports on the strength of evidence or the quality of the articles reviewed. Others noted giving greater weight to research undertaken by those with particular expertise in the field. A number attempted to "rate" papers in various ways, whether by external criteria or internal consistency:

> "We rated them, but didn't weight them in a formal way. We rated them on the relevance to the topic area and then according to whether they tried to use data … to test whatever theory they were proposing."

> "We examined the propositions put forward and the arguments to support those propositions, and then we commented on the arguments in terms of premises the argument was using or in terms of other widely accepted premises – and that comment is a kind of weighting."

Researchers on the one project that did weight studies experienced some difficulties in the process, including insufficient studies and problems with measuring external validity with the instrument used.

Secondly, with respect to the methods used to analyse and synthesise the literature, there was enormous variation from what seemed to be a highly systematic approach to what might best be described as "muddling through". It is difficult to indicate numbers here because the descriptions offered by the researchers were themselves both variable and multidimensional. Some described the process of initial analysis in great detail, but said little about how their research was synthesised and vice versa. Furthermore, it was difficult to categorise many answers. Some outlined complex analytical systems they had set up and then admitted that these had not proved particularly helpful. Others suggested they had an "*ad hoc*" approach, but then described methods that might well be viewed as systematic. There were also differences in methods, depending on whether the researchers were searching the literature to analyse the logic of arguments or to collect conclusions; those with the former approach generally seemed to find the process more straightforward.

As a broad generalisation, most researchers attempted to be reasonably systematic at the analysis stage. They tended to classify the studies collected, using hand-recorded sheets or the computer Reference Manager system. Several spoke of putting articles into piles according to their classification system and some developed tables setting out key variables. In some cases, the categories were obvious from the start; in others, they were developed in the process of building the database in the first place:

> "When I was handsearching all the journals … I was picking out those papers which were relevant and putting them on a database and, at the same time, developing a way of classifying those papers … That developed as I went along and by the time I was looking at the couple of journals at the end, I was able to classify all new papers by this system."

Some researchers found the synthesis stage equally straightforward. Using their classification systems, they were able to move directly to writing:

> "I would describe it as a descriptive synthesis … I compiled a table and put all the information together for each subject area and then, by looking at the results, tried to draw out some sort of overall conclusion."

> "We put papers into different piles and of course some papers ended up in multiple piles. And then we would try to summarise the main themes from each of these … and then that was incorporated into the report."

This was particularly easy when researchers had thought through well in advance how they would use the information collected. Some had worked out a plan for their final report at an early stage and had set up systems to collect information based around that structure. The writing-up stage then flowed fairly smoothly from that information:

> "It was a question of reading each article, identifying issues and then attempting to structure all the issues. But because … there was already a quite obvious structure, within that we just wrote up the different issues."

> "We structured it … as strongly as we could and then we tried to do a lot chapter by chapter, as implied by that structure, rather than saying 'Here are 1000 references and we've just got to plough through and put them in order' – which would have been impossible."

Indeed, one researcher who had not gone about the study in this way suggested that, with hindsight, the project team should have done so:

> "The correct strategy would have been very early on … to have identified a structure for the synthesis and then, as the papers were reviewed, the aspects of them should have been slotted into the synthesis, so that the report became written as it went along. Deep regrets on that."

Some difficulties were also encountered at the synthesis stage. In two projects, researchers found that their initial coding mechanisms proved unhelpful in practice. The sheer volume of material caused problems to others, requiring some method of determining priority of attention. Most notably, a number of researchers considered that it was only at this point they could see what their study would be able to do. They tended to argue that a synthesis could only be developed out of the information collected and could not have been preplanned on the basis of prior theory:

> "Fundamentally, it is a grounded theory approach – identifying the main themes and looking to see if they are telling you something, rather than coming in with a strong theory of this is how you should do it."

> "Because none of [the papers was] primary research – where you could extract data to synthesise and get something from the synthesis that you couldn't get from the individual papers – it became a matter of categorising, pulling together sets of arguments, and synthesising our arguments from these … It was different from a traditional systematic review."

Such researchers spoke of an "iterative" approach and several described their reports in terms of "mapping":

> "I think 'iterative' is an extremely good way to describe what was going on; we were exposing one writer's arguments to the critiques of another and so on. For us, it was an intellectual exercise in … mapping the field, in the sense that, in a highly diverse and contested area, we were bringing our professional skills to bear on sorting out frameworks within which those different positions could be understood."

> "It was very much a personal synthesis … As we were doing the review, we built up this checklist and, in effect, that helped us to develop this intellectual structure. Then that helped us to map things …"

For a number of projects, the iteration was aided by the fact that more than one researcher was involved, so that there was a process of responding to each other's draft: "It's a qualitative process of someone leading into synthesising material and having it critiqued." Several mentioned that they thought their position was stronger because they could show that it arose from a number of different people.

Some researchers appeared almost apologetic for a tendency to pull their material together on an *ad hoc* basis:

> "We had to live with a lot of gaps … I don't think it would be regarded as strictly systematic."

> "It was very patchy – hard to interpret and hard to bring together. There was no systematic or formal method of synthesising the literature, it was a bringing together and coming to some sort of view …"

Others argued strongly that such an approach was inevitable – and even desirable – for research of this kind:

> "This is more of an art than a science … It's reading, assimilating and trying to report what's there, which is a process quite different from a Cochrane style systematic review, where the whole thing can be relatively codified and done by reproducible criteria. That was not, in our view, either feasible, relevant, worthwhile or anything else for the sort of review we were writing."

> "[There is a] great desire at the moment that everything should be explicit, transparent and repeatable – it isn't like that. Theory does not come from a team of people doing transparent research; it comes from clever people sitting in garrets having clever thoughts and putting ideas together in novel ways or coming up with completely new ideas very occasionally. And if we knew that process, we could train Einsteins. The fact is we don't, so this has to be resisted to some extent."

Indeed, one interviewee spoke enthusiastically of the intellectual buzz engendered when the field was being expanded:

> "That was the part that kept the team going. If it was just 'What's been done today?', they thought that was a bit boring … But the excitement was really trying to push on a bit, because the area has basically been in the doldrums for a long, long time …"

When asked if the process of analysis and synthesis was as expected, the great majority of researchers ($n = 27$) said that it was a fairly standard approach to the kind of material being dealt with, although this was followed in some cases by a discussion of a particular difficulty that had not been anticipated. Many added that it had been more work – and more time-consuming – than expected because of the nature or amount of literature uncovered, although some said they knew at the outset that it would be difficult:

> "I didn't expect it to be that easy, but it was even more difficult because of the rather eclectic nature of the literature … It took a lot of mental effort to weave into some sort of coherent whole, without overinterpreting it."

> "I knew it was going to be awful and it was. So, no, I'm not surprised. It was predictable, completely predictable."

A few junior researchers noted that they did not have expectations, owing to their lack of experience of this kind of work.

Research methods other than literature reviews

Only five projects involved additional methods of data collection beyond the literature review (or additional analysis of data in the literature). These invariably entailed the researchers seeking to tap into the views of (non-project) colleagues in one way or another. In only two cases were these methods fairly formalised: one entailed interviews with key individuals and the other a postal questionnaire. These were used to obtain such colleagues' views on the literature and on issues beyond what had been published. Others used contacts with other researchers to help in various ways, including to carry out quality ratings of some literature, prioritise topics, discuss analysis or simply find additional literature.

Conclusions and recommendations

For a Cochrane style systematic review, prescriptive conclusions are crucial. It is from analysing the results of many trials or other forms of testing that the impact of a particular treatment is assessed. Methodological reviews, in contrast, tend to be of a different order and this was reflected in the discussions with researchers.

Some researchers (n = 6) stated with certainty that they had not come to any conclusions (or would not do so, when they wrote their report) because the concept was "problematic" in the context of methodology or the conclusions were inappropriate for this kind of study:

> "I am not sure that you can draw conclusions from methodological research … We certainly ended up being able to say these methods are not good because of this reason and these methods are good because they overcome those problems … We couldn't make conclusions in the sense of 'This is the best method, everybody should be using it.'"

> "I'm not sure that any philosopher worth their salt would say that they've ever drawn any concrete conclusions."

The majority of researchers did, however, indicate that they had drawn conclusions of one kind or another. These tended to be summaries of what was known about the subject, including, in some cases, the lack of evidence in an area. In two projects, researchers noted that they distinguished conclusions by the amount of evidence on which they were based and, in another, by the extent to which any conclusions would be affected by acquiring further evidence. Some clearly felt very constrained by the paucity of the evidence found on their subject.

All researchers said that they had made (or would make) recommendations, although some effectively conflated these with conclusions. It was not always clear who the recommendations were for, but many were clearly directed to future researchers, regarding how to design, carry out and report on trials or other research. Some were for others within the wider research enterprise, such as: managers, on how to interpret research findings; ethics committees, on how best to assess research; or journal editors, on standards of reporting research. Several researchers were keen to point out that their recommendations were fairly tentative and not of the kind found in a traditional systematic review:

> "We weren't able to say anything like 'You should always do X' because the field is too extensive and variable to be as explicit as that. But there are some principles … 'This approach is better than that approach, generally.'"

> "They are not the sort of precise recommendations that a Cochrane style review comes up with, which says that drug X is effective or surgery Y. These are policy recommendations expressed in rather general terms."

In addition, many projects specifically mentioned making recommendations for further research, directed at commissioners within the HTA Programme or beyond.

Dissemination

Although not all of the projects studied had been completed, all the researchers had given some thought to dissemination of their findings and many had already published in one location or another.

Views on the audience for their findings tended to cover four main categories: other health service researchers, national and local policy makers within the health service, research commissioners, and clinicians. Some mentioned all of these, whereas others looked to only one or two, but the frequency was in that order. A few also mentioned more specialist groups, such as ethics committees and drug companies.

Without question, most of the researchers (n = 25) believed that other researchers were their principal audience. Within this broad category, some distinguished those with a specific interest in methodological issues and some made a distinction between academic researchers and those working inside the health service. A few proposed very specific users, such as researchers about to embark on a clinical trial or those needing to interpret the results of research. It must, however, be added that some gave the impression of not having thought through very carefully how their research would be used.

In terms of actual dissemination, all researchers had plans regarding what they were seeking to do and some had more or less completed all such activity. All expected to produce several publications, including their final report in some form and, generally, contributions to two compendiums, one already published by the *British Medical Journal*[5] and the other being assembled for publication by Sage.[6] In addition, most had their sights on publication in academic journals; indeed, of the 19 projects, seven had already had articles published in such journals as the *British Medical Journal, The Lancet* and *Health Service Research and Policy.*

In addition to publications, many of the researchers had been active in presenting their findings to workshops and conferences. Some (perhaps most?) of these were academic occasions, but at least one researcher referred to contributing to a workshop for GPs. Several suggested that project results should feed into guidelines for practitioners and managers, but it was unclear how many were undertaking work towards that end. Few researchers appeared to be looking beyond academic publications.

A small number of other consequences of the work were noted. Several researchers had used the project as the basis for further research on the subject and one said that it would provide the literature review of a PhD dissertation.

Some researchers viewed any further dissemination as a job for the Department of Health or the HTA Programme itself. It was suggested that they should consider how best to disseminate the executive summaries from such projects, for instance, to ethics committees or others "at the political and administrative level" who may benefit from the results. Several suggested that the HTA Programme should put all such project reports on the Internet. Another proposal was that workshops should be set up for commissioners. Several researchers were also concerned that other funding bodies, such as the Medical Research Council, should be apprised of their results.

On the other hand, some scepticism was expressed about dissemination. This was partly about the need to avoid certain kinds of activity becoming automatic, such as placing reports on the Internet, with no thought to their utility. There were also concerns about the extent to which those apart from academics would take a real interest in the results:

> "I have conversations with NHS managers who clearly don't have either the time or the ability to read even the most widely available bits of information. So the idea that we're going to get these rather esoteric things through to them is most unlikely."

Researchers were not specifically asked about any problems associated with dissemination and few volunteered ideas here. One, however, did note the perennial difficulty of finding time for such work:

> "We've been a bit tardy in getting publications out, but … the resources are no longer there to do that – and what we're trying to do is fit it in with other projects and sort of steal a bit of time and money from those to be able to finish this one off."

The need for flexibility

From the outset, we had hypothesised that researchers undertaking this kind of study would seek to find more than one way of approaching the questions under consideration. This could be a slight amendment to the topic or the addition of new methods, but they would have a concern to improve the research by a process of triangulation. Our hypothesis was not confirmed by the data.

Changes in topic

Researchers were first asked if they had needed to change their topic during the study. Not one suggested that his or her study was significantly different from the initial protocol, with some being very firm on this issue: "No, absolutely stuck to the brief." Indeed, one respondent stressed that the proposal had been specified very tightly to avoid the problem of finding themselves with too broad a brief: "Some referees didn't like the fact that we did that, but in retrospect it was a very sensible move."

On the other hand, in 11 projects, the researchers thought that they had made modest changes to the original plan, most commonly refining the topic to fit the time or data available and, to some extent, their own sense of priorities. Other reasons for reducing the scope was that a section was being covered by another project or an initial hypothesis proved so untestable in the light of the data that it was inappropriate to pursue it. In contrast, in four projects, the topic was slightly expanded, due to the discovery of new relevant literature or to the fact that the literature covered by the project as initially specified was simply too scarce.

Changes in method

The researchers were then asked if they had changed their methods from their initial proposal. As noted above, only five projects used methods other than the literature review, in every case soliciting the views of other researchers either formally or informally. In two cases, these methods were set out in the initial protocol, although one researcher believed that there was little enthusiasm for this course from the funders. In the others, the idea came to them at some later point. In addition, in one project it was decided to add a worked example to the report as a means of clarifying what would otherwise be a very dry exposition of methods.

Whatever the problems experienced with their reviews, very few researchers considered that, with hindsight, they should have approached their task notably differently in terms of either data collection or analysis and synthesis. Most said they had gone about their study in the right way, although some added that they had lacked the time or resources to do it as well as they would have liked. Among those researchers who could see benefit in additional methods, the most common proposal was to communicate more purposively with others who had done (or were doing) the same kind of work, either early on or towards the end of the project, for instance, by circulating draft reports. One suggested that a reading group that met regularly was the obvious solution to keeping in touch here.

A few added project-specific and practical ideas about what they might have done with hindsight. Many of these revolved around the search process, such as including more databases, doing more hand or electronic searching, making more use of the Internet, clarifying criteria for inclusion or exclusion early on, and keeping better records on information collected. Some involved the later stages, such as double-checking statistics, weighting studies or using a worked example; others involved the early drafting of a skeleton report.

Only one researcher considered that the project would have benefited from an entirely new kind of method, in addition to the literature search undertaken. The proposed addition was a survey of people involved in randomised controlled trials, seen to be needed because the topic addressed was one that such people would know a lot about, but on which little had been written.

Views on changes in topic or method

Finally, researchers were asked about the appropriateness of a change in topic or method once a project was under way, in terms of validity and the HTA Methodology Programme. Somewhat over half (n = 19) argued clearly that a change of topic would be problematic for the funders and would be unreasonable unless it was carefully checked with them:

> "If you change your topic, you're moving the goal posts entirely – so what you get at the end of the day is going to be substantially different from what the HTA Programme has been led to expect."

Several participants spoke of such an action as breaking a contractual obligation. One added that, with a 1-year project, it would be difficult to accomplish much if the topic were changed once the study had started.

However, almost as many researchers, including some of the above, could see the potential need to redefine questions in the light of information gained or thinking undertaken at an early stage. What was important was to have some system for dialogue with funders when a topic was found to be irrelevant or unfeasible:

> "If you ask a senior group to take on something, they will naturally and inevitably refine that topic in some way or another, either narrowing it down or redefining it. That's just the way the world is; they are going to give a subject more attention than the commissioning process can."

> "The commissioning group … should recognise that [this] may quite reasonably happen in some cases, that it's then a matter of negotiation … Everyone's got to be honest and not feel that they've got to pretend they're doing something they're not doing because there'll be trouble if they don't."

Researchers were more reluctant to comment on the impact on validity of a change in topic, but those who did so tended to argue that it was very difficult to generalise. Validity related to the aims and methods for a topic, so it depended on both the new topic and new methods:

> "Validity needs to be judged against what you set out to do. So if you are explicit about what you are doing, what you are not doing … then the judgement of validity should be made against that explicit statement of the aims and objectives."

> "Validity is tricky to generalise. It depends whether people acknowledge the restrictions of their topic. If they're very explicit about what they're trying to say or why they're saying it, then it would be valid on its own terms. The problem is if they're pretending to do one thing and are in fact doing another."

With respect to changes in research methods, most respondents did not see this as a problem for the HTA Programme or for the validity of a study, provided there were good reasons for the change. Indeed, several argued the benefits of some flexibility because changes in method (even more than changes in topic) were often found to be necessary after some initial thinking on a project:

> "I don't see that should be a problem for the HTA as long as they're getting what they are wanting an answer to. And I would imagine that mostly the change of methodology will be to improve validity rather than otherwise."

> "If there was a means of consulting on that, getting agreement – in all of these cases, provided everybody's happy, it's fine. It's wrong to stick rigidly to something if it rapidly becomes clear that it's inappropriate."

A few researchers did express concern about methodological changes, however: "You really should think through the methods beforehand." A number argued that any changes in method should not be on a large scale:

> "If you said you were going to do a systematic review and then you ended up just getting a few things that you happened to have, then that would be a problem."

Furthermore, several stressed the importance of documenting and reporting changes, so that future researchers could learn from them. Again, there was some concern about the potential for change

in a very short-term project. Piloting was suggested when there was doubt about methods.

The question of bias

One issue of some concern from the outset was the extent to which systematic reviews in this context could be objective or unbiased. Although a case can be made that there is no value-free research in any context, the kind of qualitative review undertaken by most of the researchers studied seemed particularly problematic. They were therefore asked whether they thought it was possible to achieve a value-free presentation of results and, if not, what they had done to minimise the problem.

Very few researchers ($n = 3$) stated without qualification that a value-free presentation was possible. Considerably more ($n = 10$) thought that it could be achieved – or more easily achieved – in certain circumstances, for instance, with respect to quantitative results:

> "I look at it as a bit of a sliding scale … You can be pretty objective on new statistical methods for doing things – it is written in algebra, it's there … It gets a lot more difficult when you are working with more qualitative ideas."

> "You can do bad research in any field … In quantitative work, people … can misrepresent data to their own ends very easily. But I think it is easier for the honest researcher to be honest about the pitfalls in quantitative research than it is in qualitative research."

The majority of researchers ($n = 22$) argued that it was impossible to achieve a completely value-free presentation in the kind of reviews they were undertaking. This was partly because of the need to assess the research involved and it was very difficult to be completely objective, however hard one tried:

> "It is fairly hard, because we do tend to form opinions and that's what we end up presenting. It's basically based on the literature – our views about the appropriateness or not of somebody else's conclusions."

> "Almost every document reflects in small, subtle and sometimes important ways the values and priorities and world view of its authors. I don't think any document can ever be completely value free."

A few positively revelled in their subjectivity:

> "I'm completely biased. I wouldn't have done it, if I hadn't believed in the stuff … But I think it's possible to state your biases."

Many researchers believed, in the words of one, that objectivity was "something one should strive towards". Certainly, the great majority ($n = 26$) considered that they had themselves made specific efforts to overcome any inherent bias. Their avenues to this end differed considerably, however. Some argued that the real issue was to present data as clearly as possible, suggesting that they had tried to "report on what we found" or "let the data percolate up". As one explained:

> "We had a structured way of getting data from the papers and we then presented that information as it came out. We didn't do any further analysis on it and then we tried to summarise it. And … for a particular recommendation, we said the number of studies contributing to that. So we tried to make it that the reader can make their own judgement about whether a particular approach would work for them, based on the evidence that we found."

Separating findings and discussion was also seen as part of a similar process:

> "In our particular area you can set out a value-free description of the issues … You have to then move forward and take a value judgement about what matters and what doesn't and, from there on, I would have no qualms about being told my approach was subjective."

Another approach, in contrast, was an effort to be inclusive of all points of view, giving sufficient voice to differing positions:

> "We made a conscious commitment to identifying and presenting all positions on any issue. Where we were obliged to come down on one side of the argument, we made explicit that this was our view, having weighed the arguments, but that other scholars might reasonably have come to different conclusions."

Many researchers also spoke of their efforts to be explicit in their assumptions, values or the basis of their recommendations:

> "What you can do is try to be explicit about your position and as respectful as you can of other people's position. And try to present them in a way in which – accuracy, of course, is not always even attemptable – they would at least agree with it if you put it to them."

> "The most important thing is for me to think very carefully about the potential pitfalls – what are the opportunities for bias and so on? And point those out to the reader, which I do whenever I write anything up."

A few recognised that they were unlikely to be wholly successful in this regard, for several reasons:

> "In the tradition of scientific and social science writing, you're brought up in a writing tradition that bends over backwards to sound impartial, when in fact you are pushing your own point of view like crazy … It is probably better just to acknowledge up front that this is your intellectual background … "

> "Even with every will in the world to be objective, if you have read a really well-written paper pushing one point and the paper pushing the other point isn't as well written and doesn't put all the things forward, then … you are going to be biased towards the better written paper … And if there are five papers against and one for, doesn't it mean that the against are more valid?"

Yet another device employed to reduce bias was to test out ideas and analysis on other people, including those known to have differing views. In some cases this occurred naturally (or purposively) within the project team; in others, new people were brought into the process for exactly this reason:

> "[We aimed] to get ten senior people, not in our pockets metaphorically speaking, to tell us exactly what they thought of the draft and try as honestly as we can to revise it in the light of their views."

The researchers whose projects had been refereed were asked if they considered that the peer review process had been helpful in making their results less value laden. Most (16/27) thought that it had made no difference in their case, although many of them believed it could make a difference in other circumstances. A variety of reasons were given here. Several argued that referees became involved too late in the process for any major changes in response to their comments: "Big overhauls can't be done, you can't start the project again." Indeed, it was suggested that earlier attention to this issue by the initial referees could be helpful. Others suggested that their referees had not made any comments relevant to this aspect. It was also argued that the referees were themselves researchers and therefore no more objective than anyone else.

Some, however, did consider that their referees had been helpful in reducing bias in their report. They had called attention to relevant issues and forced them to be more explicit: "It made us think a bit harder about certain statements."

Finally, a few argued that bias was not really an issue in their case; they were genuinely open minded or value free:

> "We were aiming to be neutral the whole way through. I can't really describe how we did it, but we set out with: 'We don't know, we have no preconceptions either way,' and we just wanted to try and dispel some myths."

One researcher noted that the issue of subjectivity was much less important in the methodological context:

> "When you are trying to review all the studies that have tested a treatment on a disease, any biases involved could end up with you giving the wrong conclusions about the treatment. But biases here … are much less important – just maybe you place more emphasis in describing a method than you would."

Project management

In addition to initial concerns about the difficulties of undertaking systematic reviews in this field, the respondents were also concerned to address questions relating to project management. It was their view from the outset that it would be particularly important for these projects to be closely managed by the principal investigator(s).

Timing and time management

None of the projects met the original deadline. The length of their overrun was not specifically asked in the interview but, where researchers volunteered the information, it was clear that it was from a few weeks to nearly 2 years. In two cases, principal investigators claimed to have met the deadline, but were contradicted by the junior researchers as having been 3 months late.

Ten projects (out of 19) were completed within 3 months of their due date. The researchers tended not to consider this as a problem; indeed, a number referred to a 3-month grace period, which they understood to require no further explanation. The principal investigators tended to explain their relative promptness as arising from the "quality" or "efficiency" of those doing the day-to-day research. The researchers, conversely, considered they were well managed, with clear guidelines about deadlines for the different stages.

Those involved in three of the projects admitted to being late by 12 months or more. The reasons given for delayed submission will be familiar to anyone involved in managing researchers. Over and over again, researchers talked about the amount of work to be done, complicated by the fact that some principal investigators were involved in more than the one project:

> "The grantholders were, fortunately or unfortunately, all … very senior and it was a problem for them to get it to the top of their in-trays."

> "If it was delayed, it was by the sheer bulk of work. On reflection, the scope of the review was probably larger than was wise under the circumstances and that created certain difficulties."

These explanations, however, are little more than a rewording of the problem.

It became evident that the real problem was one of time management, allocating time between the different stages of the research process. Researchers in only six projects considered that they did not have any substantial time management problems. Even among these, some thought that some aspects of the work had taken slightly longer than anticipated or that, in hindsight, they might have allocated their time somewhat better. They tended, however, to argue that they were well organised from the beginning, with clear allocation of responsibilities across researchers and good team working.

For the others, four stages seemed particularly problematic.

First, there was the time taken in collecting and reading an extensive literature, often very much larger than had been anticipated. Some of the problems were very practical:

> "One of the biggest time things is actually getting the references on your desk in some form – filling in library request forms, chasing them up, things out of libraries, things wrongly cited …. "

One researcher also recounted considerable practical problems arising from a limit on the number of interlibrary loans allowed per day. Indeed, some had not anticipated that they would have to pay for such loans, placing additional pressures on their budget.

Others problems represented intellectual or even emotional difficulties:

> "The main problem was knowing when to stop looking for things, because there was certainly a tendency to just keep looking … "

> "The problem with handsearching is because it's so tedious, you tend to get very bored very quickly – you have to vary what you're doing."

> "It got a bit lonely. I turned into a bit of a recluse and a bit withdrawn at times, sitting there reading 20 papers a day and writing notes."

Secondly, there were problems at the analysis and synthesis stage, which was often delayed because of the length of the search period. Some researchers found it difficult to get started on this part of the work. In some projects, the delay meant that this stage was undertaken by more senior staff, who found it difficult to find the time:

> "We didn't quite know what we had let ourselves in for. We got a bit obsessed with the search – and when we got everything, we probably spent far too long on that and shied away from the really difficult task, which was how on earth are we going to make sense of all this?"

> "I took the lead in overseeing the project and … ultimately in synthesising the literature. And, because I'm a very busy person, that was very difficult to timetable, very difficult indeed to find the dedicated time to actually sit down and do that."

Thirdly, there were problems with finding time for writing. This was particularly problematic for those whose day-to-day researcher had moved on:

> "I certainly underestimated the writing-up phase. I should have kept a much more careful eye on [the research assistant] and made sure that he was writing stuff that was coherent. But it was difficult for him to write clearly when the exact way in which it was all coming together wasn't clear until very close to the end."

> "This in effect involves understanding methodological developments and writing coherently about them in a very structured way. That's quite a demanding task in terms of time and expertise in itself, and it covered such an extensive range of literature. You did it for one chapter and then you've got another 20 chapters to do it for – there was a lot of work involved."

Problems also arose when teams of researchers were commenting on each other's drafts because this entailed "an awful lot of toing and froing" at the end of the project:

> "Eventually you have to just say 'stop', because you are recorrecting what someone else has recorrected. I had to take the lead on that … because it could have gone on for ever."

Finally, in some cases, a delay arose at the later stages owing to referees. The projects necessarily had a period "on hold" – which was often quite prolonged – when the authors were waiting for referees' comments and then some time was needed to respond to them. This was a source of considerable annoyance to the senior researchers involved, especially when junior research staff – and occasionally the senior ones themselves – had moved on: "It was just an absolute nightmare." It is clear that many researchers worked very hard to make up for the time lost, devoting weekends and evenings to getting their reports finished:

> "It took a long while before we began to realise what it was we wanted to do and get cracking. Once we did, we worked like the clappers for quite a long period. We really had to struggle to produce the report within the original timetable."

Many junior researchers thought that they had learned a lot about managing their own time and keeping fellow workers abreast of their own work and thinking. Again and again they noted that they needed to break down a project into its separate components and plan their time carefully and realistically. Several mentioned that it would have helped to have some training or, at least, some early contact with others who were undertaking similar work.

Day-to-day management

The question arose of the extent to which there was good research management by the senior staff involved. Of the 19 projects, 13 were clearly considered to be either well or reasonably managed by both the senior and junior staff involved (although, for two, only the principal investigator was interviewed and, for one, only the junior researcher was interviewed). In all these cases, the respective roles of the senior and more junior researchers were clearly specified and the latter were actively supervised.

There were substantial differences between these projects in the degree of responsibility given to junior researchers and the relative time input of junior and senior staff, but the key factor seems to be clarity and careful monitoring. In every case there were fairly frequent meetings, ranging from weekly to monthly, or very occasionally more frequently; in one project, these meetings lasted a full day. Researchers also noted a sense of easy availability through day-to-day proximity; "a lot of stuff [was] done in the corridor informally" and via telephone or e-mail.

Researchers also stressed the importance of establishing boundaries and a set timetable. In some cases these were reassessed over the course of the project, but in others they were determined very early on:

> "In your original proposal, you always say something about what you are going to do, but in the first two meetings also we had some discussions about what we would expect to get out and what shape and when."

There was also a need to set clear deadlines for drafts, as well as for comments from colleagues. Another key decision was to start writing as early as possible and to give timely papers to conferences to force early thinking on the issues at hand and to obtain prompt feedback on external reactions to that thinking.

In three projects, both senior and junior researchers clearly recognised that project management had been a problem. Although they started reasonably well, with regular meetings, these were less closely managed and slippage became a problem. This was particularly keenly felt by those at the day-to-day end:

> "We had a fairly informal time line written into the grant proposal and, after about 3 or 4 months, my impression was that it had gone out of the window."

> "It was pretty hands-off. The person who actually got the grant did virtually nothing and the person who was nominated to run the project … was quite difficult in terms of finding time to manage the project … Things drifted and people didn't really know what they were supposed to be doing."

In three further projects, there was a difference of view between senior and junior researchers. In one, the junior researcher considered that the project had been well managed, but the principal investigator was more self-critical. In the other two, more expectedly, the views were the other way around. In all cases, there were differences of view on the extent of the actual involvement of the senior person, possibly arising from how much direction was thought to be needed:

> "We had some planning meetings, but when it really came to search strategies and all that, we didn't have much contact."

In some cases (and even in one or two that were considered to be well managed) logistics were a major barrier. Projects were either set up to be multisited or key researchers moved during the progress of the work, so communication was forced to be at a distance. All the researchers involved in such projects thought that this was a fairly serious problem:

> "The grantholder was doing a very good job, but it's always so much easier to manage a project when you're there. Getting me there proved difficult, because it takes quite a long time on the train. And my boss, having to get up to speed in a new job, also had other things going on."

The question arises of whether those projects that were well managed were also completed closer to time than the others. Regrettably, as noted above, information was not collected systematically on the duration of overspill. Nonetheless, where data were given, three of the six projects said to be poorly managed were

also the most delayed in completion (a year or more) and all of those said to be well managed were completed within 3 months of the deadline.

The researchers said that they had learned many lessons about project management. Both senior and junior staff commented on the need for careful time management, including early writing-up: "Double the length of time you think it is going to take."

> "We got a bit lost in searching for literature and then, when it came to synthesis, we were completely stunned how much stuff we had. So we would do it differently and focus on what would we want to write up and how much literature do we need to support it."

> "You have to be pragmatic, you have to say there is a trade-off between the ultimate in perfection and the cost of doing it and the time and all of the rest of it … There's a danger of becoming a bit like train spotting – you're doing things for the sake of doing them, rather than actually finding the answer."

Many interviewees believed that there was a need for senior researchers to take their research management role seriously, giving reasonable time to those newly working in the field. This also means ensuring that the research team includes those with appropriate skills. A few researchers, both senior and junior, suggested that it would have helped to have clear project milestones so that, even if they were not met, one could consider why not and the implications for the remainder of the project:

> "You have to lay out a very tight programme of work, with benchmarks that you are going to stick to and a tight programme of writing."

There were also concerns about clarity regarding the nature of the task:

> "Calling this type of study a 'systematic review' is actually quite harmful, because it creates the impression in some people's minds … that they're obliged to spend an awful lot of time on the actual collecting of information … "

> "Because you're dealing with theoretical literature, ideas of 'exhaustive' or 'comprehensive' aren't relevant – that can throw people into panic and doom and gloom of 'How can we do this?' They're labouring under the misapprehension that a systematic review can only be done one way."

Several researchers stressed the need to build in a sufficient amount of a senior person's time, even if this has to be paid for, because they need to do the difficult interpretative tasks:

> "The really tough part of it is not searching – it's actually trying to make sense of it. If you can just buy in a small amount of someone else's relative experience, time and attention, then that can help probably more effectively."

Other issues raised included clarifying the boundaries of the task more carefully from the outset, avoiding collaboration at a distance, and exercising much tighter project management. Of course, there was also a sense that more resources were needed to do the job properly:

> "Researchers are too tight with their budgets – they are always skimping and saving and at the end of the day they just burn themselves out."

In looking back over their problems as a whole, the researchers raised a number of different issues. A few projects clearly had staffing problems, both in terms of unanticipated departures and in researchers' suitability for the work. Two principal investigators considered that they had been unable to hire junior researchers with appropriate qualifications. Moreover, the process of involving many people meant, as one person noted: "The project moves as fast as the slowest person."

Some argued that their project was substantially under-resourced, which was particularly strongly felt when they had been asked to take on an additional aspect of work or if their initial budget had been cut down. One researcher was particularly angry at the dilemma the project thus faced:

> "We had accepted the contract because … they said 'yes' in principle and you can assume that you are going to get it and that there would be a little bit of a negotiation about detail. [That] turned out to be them telling us that we are going to do about 50% more work within the [same] time and budget … The university would have been furious if we had just sent back the contract."

Many researchers were, however, willing to admit that the job was simply more complex than they had anticipated. There had been a degree of overoptimism, a lack of foresight, combined with poor management, which inevitably led to delays in completion. This kind of review was simply more complex than others. One researcher summed up the problem rather clearly:

> "It was such a wordy subject; there was such an enormous volume of words. It wasn't like a scientific thing which is reduced to figures. A huge volume of actual words to be read and assembled and categorised and written up."

And one was lost for words altogether:

> "Um … Ah … Oh dear. This is um … I don't … Oh dear. Oh. I don't know if I … Can I not answer that one, please?"

To end this section on a salutary note, eight researchers (more than one-fifth) said either that they would never do such a project again or they would advise others not to do it, although the question of their willingness to do it again had never been explicitly asked.

Involvement of others

Many of the projects had a number of grant-holders. Although all of these people were fully involved in some projects, this was not invariably the case. When grantholders were not especially active, most researchers were sanguine about the extent of their involvement, noting that their role had been expected to be different from the start. They had been put on the initial application to offer specific technical expertise when needed or, quite simply, to give it "credibility". The individuals tended, at a minimum, to comment on drafts produced either during or at the end of the project.

Only a few researchers suggested that their grantholders could have made more of an effort to attend meetings or provide comments. One or two felt quite angry at the way they had taken credit without doing much work:

> "People can get their names on things without having to do anything, except be important … I think if you're going to get your name on a grant, you ought to be contributing some concrete thing. If you're going to take the credit, do something to earn it."

Those researchers who thought that they would have benefited from greater involvement from senior people were asked if some specific funding for this purpose was a good idea. The answers here were evenly divided. Some considered that such an incentive would get such work "to the top of the in-tray, rather than half way down". Others were concerned about the extra costs it would entail or argued that money either was not – or should not be – the issue:

> "In this sort of study, the expertise will be provided freely or it won't be provided at all, in terms of input from experts in science and medicine."

In two projects, such funding had been expressly built into the proposal to ensure a reasonable level of activity. It was considered that it had accomplished this purpose, challenging the researchers' thinking and thereby making their projects stronger. It was thought that they obtained a better level of involvement because of the funding incentive.

Only five of the 19 projects had some form of steering or advisory group beyond the grant-holders. These were used in some cases to bring in expertise from other fields and, in others, to provide a sounding board where criticisms could be voiced and arguments vetted. They were generally viewed to be helpful, although some found it difficult to bring them all together because of busy timetables. In one case it was also difficult when they were brought together because of conflicting points of view. In addition, a few other projects circulated drafts to other people for comment.

Most researchers (n = 19) agreed that, had they experienced problems in keeping to their initial protocol, some kind of formalised steering group might have been helpful to serve as a sounding board and to offer advice, as well as to offer "moral authority" to amend early decisions:

> "Civil servants pay a certain amount of respect to academics and let them do things their own way, but then they grouse about what they get. So these collaborations work better if everyone works as a team, including the commissioners."

Project managers tended to see such benefits more than the junior researchers. Several suggested that it would be more helpful to have an individual to contact, rather than a whole group, as discussions could then be held, particularly with commissioners, on an *ad hoc* basis.

Nine researchers thought that this would not have been helpful. They tended to argue that they knew where to go for help if it was needed and that it might prove "just another meeting". There was a certain wariness about steering groups, even among those who deemed them as possibly helpful, because they could sometimes be obstructive with little subsequent improvement to the quality of the research.

With respect to who should comprise such a steering group, most researchers tended to make general statements about "someone from the commissioning group" or "other experts in the field", although a few named specific individuals who might have been helpful. Some said they should not be administrators from the Department of Health, on the grounds that what was needed was more academic expertise. On the other hand, some argued that someone from the HTA Programme would be best placed to understand what the issues were.

Most researchers (n = 24) believed that the HTA Programme itself had exercised an appropriate

level of involvement; only seven considered it had not done so (four others did not know). Principal investigators were especially likely to think that the Programme had been sufficiently involved. Researchers tended to argue that the HTA Programme had not done very much, but they could not see what else might have been done. It was quite right that they were given a grant and allowed to get on with it, as long as no major changes were introduced. Several commented that they felt able to ask for help. One explicitly noted that his or her project had been appropriately hassled to get the report in on time.

Those who considered that they would have liked more involvement from the HTA Programme noted a need for more steering of one kind or another. One principal investigator was particularly critical:

> "I wouldn't have commissioned us to do what we said we were going to do in the proposal, because I couldn't have worked out what that was. It would have been helpful all round if there had been more focused thinking about what it was that they really wanted to know."

Another researcher suggested that the project would have benefited from some support when their referees' comments criticised them for not doing particular work that they had never been asked to do. Several, including some who thought the involvement was appropriate, noted that they would have benefited from milestones for their project and/or some reminder when interim reports were due. A few also commented on the need for better chasing of referees' reports because this caused considerable problems at the later stages of the project.

Collaboration

Of the 19 projects, six were clearly seen as having little or no overlap with any other HTA-funded project. The 11 others were believed to have some overlap with one or more other projects; some form of collaboration was generally developed in consequence. This sometimes involved one or more joint meetings (over and above the more general Methodology Group meetings, discussed below) to iron out how best to avoid duplication, discuss methods and references, and share problems. In addition, there was often considerable telephone discussion and much correspondence, including the exchange of early drafts to use each other as sounding boards: "It was a sort-of 'Show me yours and I'll show you mine.'"

Two sets of researchers, in contrast, decided deliberately not to collaborate too closely in order to develop independent replication of roughly the same brief. In both cases, there was some contact between them and a willingness to keep the other project informed about references, but a certain distance was nonetheless maintained. A few researchers expressed uncertainty about the extent to which they were expected – or allowed – to collaborate. A number also raised broader issues concerning the relationships between academics in the same business:

> "They were almost rival projects and there was the question of do you collaborate with your rivals? Does the commissioning body want two independent views of the same thing? That's an area in which clarity of guidance from the commissioning group would have been very helpful. I think they didn't know … whether they were after replication from two independent reviewers or whether they wanted collaboration."

> "There's a lot of competition in research, as I'm sure there is in other fields. I'd like to think that talking to other people would help and we're all in it to find the actual truth and what's out there. But I think the competition is too great in some cases, so it's probably best that people work alone."

There was much less collaboration with groups outside the HTA Programme. Researchers on only seven projects noted any such collaboration and this was often little more than keeping in touch with others doing similar work. This tended to involve some sharing of ideas, references and circulating drafts and so forth. Several researchers mentioned that they had received advice on systematic reviews from Cochrane groups.

Collaboration was, not surprisingly, widely seen to have been helpful. At a minimum, it resulted in saved time in the acquisition of useful information, for instance about references. In some cases, researchers sharpened up their ideas and gained a better understanding of what they were doing. Two noted that it served as a source of moral support. On the other hand, a few believed that it made very little difference to their project.

Virtually all (n = 28) of the researchers interviewed had attended the Methodology Group meetings at least once, and generally much more often. Those who did not were either too far away (e.g. overseas) or were members of a research team that was already represented at these meetings.

The Methodology Group was widely seen to be extremely helpful, for a variety of reasons. The researchers welcomed the chance to make contact with others carrying out similar work, exchanging ideas, papers and methodological tips. Several

mentioned that it had the useful outcome of some joint publications, but, without doubt, the biggest bonus was described many times as "reassurance" that the difficulties they were experiencing were shared by others. The Group became, in their own words, a form of mutual support, particularly important in the light of the considerable problems the researchers had encountered:

> "A definite feeling that you weren't alone. You're coming out with all these problems and then you go to one of these meetings and they've had exactly the same problem – and it was great."

There was, indeed, some interest in rolling financial support to ensure that the Methodology Group continued over time, to help future researchers in the field.

Only two negative issues were raised concerning the Group. One, arising from researchers who had to travel a long distance, concerned the practical difficulties of arriving late (and tired) and needing to leave early. It was questioned whether the location of the meetings could be varied. The other was a sense of feeling excluded from an "inner group" or "clique".

Finally, researchers were asked if it would have been helpful for collaboration to have been more actively coordinated. Slightly over half (n = 19) said they thought the amount of help given was appropriate, either because there was no one else they might have collaborated with in any case or because a reasonable amount of help was already given. One respondent considered that such collaboration was much more fruitful if developed by academics themselves.

Among those who thought that additional help would be welcome there were essentially three broad arguments. First, it was suggested that more central direction would have been useful:

> "A programme director could have been more explicit about how we should divide up the cake, instead of having to negotiate this rather awkwardly with other researchers … "

Secondly, it was considered that more assistance could have been given to those who were working in a similar patch:

> "It does seem absolutely ludicrous that 21 of us or whatever round the country were trundling away doing similar things … For instance, we were all building up huge databases … [and] it was expensive getting these papers. If there had been some way of finding out what people were holding, it would have made life a lot easier."

Thirdly, arising from many of those who were distantly located from other projects, there was a need for more financial help to enable them to attend the Methodology Group meetings or to attend more local gatherings (or to keep in touch by other means).

Referees

Six projects had either not yet gone to referees or were currently with them. Of the remaining 13, the majority of researchers (19/26) found their comments helpful, either completely or in part. Indeed, only four researchers thought that the referees' comments were not helpful at all.

The reasons given for why referees' comments were helpful varied considerably. Some, particularly junior researchers, clearly welcomed their good report:

> "Five of the six referees were very positive … I was over the moon for days afterwards, because it was my first piece of refereed work."

> "They were very reassuring, so they were helpful in that sense … They didn't really suggest too many big changes. It was certainly nice to have those – if we just handed it in with no feedback, that would have been pretty dreadful."

Others welcomed negative comments because they made their report stronger in the end. This included attention to the structure of the report, as well as the arguments addressed:

> "It was very critical, unfairly critical, but actually that was very helpful … It enabled me to phrase the argument even more tightly so that any wilful or incidental misreading is less likely to occur than otherwise would have been the case."

Those who found the referees' comments unhelpful (including some who had mixed views) said that they were being criticised either for not doing work they had never intended to do or for not following the procedures for a proper systematic review:

> "They got hung up on at great length that it wasn't a Cochrane type review, which … showed a complete lack of understanding of reviewing theoretical literature … They were thinking of it entirely as widget research, as to whether one widget is better than another. But it's not like that when you're dealing with conceptual literature."

> "Both reviewers commented on things we hadn't done, stressing what we'd missed, but in our original protocol we never said we would do these things … That was slightly disappointing – that they were so negative in that sense."

Several respondents found the comments to be footling, for instance concerning typographical errors and one had been told that there were inaccurate citations, without any indication of which ones.

Almost all researchers agreed that the refereeing process had been helpful in identifying gaps. A few thought that these were serious additions but, for most, they were seen to be fairly minor.

The majority of researchers whose projects had been refereed considered that their referees had the right skills, although some thought not all of them did. Many felt unable to comment on this question. Such skills were deemed by most to be a mix of specific professional and methodological expertise, such as, for example, statistics, epidemiology, economics or philosophy, depending on the project. In addition, a considerable number thought that referees should have experience of carrying out a systematic review. Several suggested that at least one should have a broader experience within the NHS. A few others proposed that at least one reviewer should be an intelligent generalist, partly to check that the report was comprehensible to someone outside the field:

> "If you write something that only a health economist could understand, then there's not really much point writing it, is there?"

> "[They] need to have reasonable skills in analytical and critical thought … in other words, somebody with both their own professional expertise and an ability to step back and view things from outside of their own particular discipline."

The researchers were asked how many referees should have been involved in their kind of project. Among those who had a clear opinion on the issue, 14 proposed two or three, 11 proposed four or five, and two proposed six. Those who felt the need for four or more tended to be associated with projects covering several topics or disciplines, thereby requiring more viewpoints at this stage. Several noted that the greater the number, the greater the potential for delay. Two stressed that the real issue was the quality of the refereeing and obtaining people with a real understanding of the field and a willingness to look carefully at the reports. Several also suggested that, ideally, there should be some sort of coordinating process across the different referee reports.

Finally, two researchers added comments about the refereeing process. One considered that it should be more open, rather than anonymous, preferably on both sides. This would make referees more "civil" to each other, rather than having the protection of anonymity, which "brings out the worst in people". The other was concerned about the time referees seem to take. The delay in receiving comments had meant that a report could not be put in the public domain, despite considerable early interest in its content. The case was made for a small fee to be paid to help referees to complete their reports on time:

> "Not because you need the money exactly, but it just focuses you on it a bit more."

Issues for the HTA Programme

At the end of the interview, each researcher was asked a set of questions concerning the HTA Programme, around the broad issue of future research directions.

Recommendations for future research

Every completed research project covered by the researchers interviewed made recommendations for further research, although few felt able to argue that their own ideas should be seen as a high priority: "I wouldn't be so presumptuous." The recommendations generally related closely to the research that they had just undertaken; they are not outlined here because they are detailed in the reports and are, in any case, most appropriately seen in the context of the particular findings.

Here, and elsewhere, several researchers raised the potential for methodological questions to be bolted on to other studies:

> "If people that were doing a research project anyway … [had] some additional funding to build in a methodological aspect – that would be a good idea. Rather than funding specifically methodological projects, which can often be in a vacuum and then are always a bit false."

Updating methodological reviews

Virtually all the researchers thought that methodological reviews of the kind they had undertaken should be updated, but many quickly added caveats concerning the need for balance with other priorities and attention to the extent to which the particular area was very fast moving. The content of some reviews, for instance, may not change very quickly and would not need frequent updating, whereas some areas were in a state of flux and would strongly benefit from being updated in a few years' or even months' time. Several researchers referred to the need to work from an electronic database, which could be very

expensive to maintain. Many thought that an update would be relatively easy to do.

Among those who considered who should undertake such updates, it was generally considered that the researchers who had done the initial review should be given first refusal. They would do such work more cheaply and efficiently than anyone else. At a minimum, it was suggested, they should be involved through some joint meetings, to draw on what one researcher called "folk knowledge" in the field. On the other hand, some could see the case for a fresh perspective on the issues. Several researchers noted that such updates needed to be properly funded and could not be seen as part of the standard remit of the original review, unlike those undertaken in the Cochrane tradition.

Future directions for the Methodology Programme

With respect to future activity by the Methodology Programme, the researchers were asked whether more progress was likely to be made in methodology by funding more large (3-year) studies or small (6-month) scoping projects. Opinion was divided fairly evenly, with ten urging more large projects, seven proposing more small projects, and 13 saying either both or that it depends on the nature of the questions to be asked. Some also added some intermediate time, such as a year, as likely to be most appropriate. An additional five respondents did not know or did not answer the question. The preference for large projects was particularly notable among principal investigators; indeed, nine of the ten votes for larger studies came from them.

The case for large studies rested on the arguments that some issues simply took a long time to investigate and that it could take time to clarify what is needed:

> "The time scale that's needed to clarify your own thought processes – and to come from what you might put into a grant application to what you actually use as your methodology – probably needs to be a little longer than most people give it credit for. The period of a grant needs to be long enough to encompass that, and 6-month studies are not."

Several researchers also pointed out the practical problems of small projects because it is very difficult to hire a good researcher for 6 months only. Senior researchers could end up spending a lot of time trying to obtain grants, rather than getting on with the work. Indeed, one or two argued the need for Programme funding or for large multicentre projects to be able to answer effectiveness questions properly. One of these noted the problem of studies being funded prematurely and that a more iterative approach was best.

The case for small studies tended to be based on a belief that much could be learned in a short time and, in some views, considerable sums were being wasted: "By the time these great lumbering projects are finished, everyone's lost interest and moved on to other things …" At a minimum, early on in the planning process, very close attention should be given to how larger studies were going to be carried out, in order to avoid wasted resources. Here again, several researchers stressed the importance of those commissioning research knowing exactly what it was that they wanted.

The researchers were similarly asked whether more progress would be made by more primary studies (involving the collection of new data) or more secondary ones. Here, they were more reluctant to opt for one or the other: seven proposed the need for more primary research (of whom four were principal investigators) and three for more secondary research (of whom two were principal investigators); of the remainder, 16 suggested both or that it would depend on the question being asked, and nine either did not know or effectively did not answer the question.

The case for more primary research rested on the argument that a considerable amount of secondary research had now been funded ("reviewing each other until the cows come home") and it was appropriate to fill the gaps thereby uncovered. Some important questions were seen to remain unanswered:

> "People always talk about the randomised controlled trial as the golden standard of clinical research and then there's supposed to be a hierarchy of other methods, but where's the evidence for this? … Why and how much worth are these other studies? And when we are obliged to use something that isn't the golden standard, what price are we paying for that?"

Indeed, the principal case for more secondary research was for it to serve either as a source of information on which subsequent primary studies could be based or as a means of determining that no further primary research was needed.

Several researchers also referred to the pros and cons of bolting methodological studies on to other research, such as clinical trials. Some thought this was an important way to get new work done, but others were concerned that the methodological issues would not be seen as a priority and could

potentially be overtaken by concern with the content of the studies.

Finally, the researchers were asked to discuss the most appropriate methods for methodology research. Seven did not really answer this question, albeit providing some additional pertinent thoughts. Of the remainder, the majority (n = 15) replied emphatically that it depends on the questions being asked:

> "The best method is the one which is best equipped to answer the questions which are specific to the methodology in question and the sources upon which such answers can be based. It really is impossible to provide a template for all such studies – this is why properly conducted scoping studies are important."

> "Tackling the methodological area is no different to doing a good quality randomised controlled trial or a good quality case control study … You basically need appropriate research questions, appropriate aims and objectives already developed … You just can't name a single method other than you use the same methods as you use to do any other good quality research."

Among the others, three stressed the importance of a combination of methods and nine proposed further reviews, whether properly systematic or simply comprehensive:

> "Systematic reviews always have their place in a big programme, because if you haven't got that, you don't know where to go next. In any field, you need that as a starting point, after which various sorts of issues will become clear and would be taken up. So they are essential, particularly early on and updated periodically."

A case was also made for giving researchers greater scope to determine the appropriate method in the course of their projects:

> "I think you need a certain amount of funding that gives people the opportunity to do original work – for which, by definition, you can't guarantee the outcome. You can't really write a project that says, 'We will do this and that in the next thing', because you may well find that you have to approach the problem or a different problem in a different spirit. So some of it needs to be fairly free and based on trust, the trust that the majority of researchers will do something useful."

Somewhat less than half of the researchers (n = 15) believed that the HTA Programme had covered the "big issues" and another 14 said that they did not know (or did not respond to the question). Among the six who thought that there were major issues still to be addressed (and some of the "don't knows" who were uncertain about what had been funded), a few additional issues were raised. Two people proposed the need for attention to informed consent, especially with populations other than competent adults. There were also concerns about the cost of randomised trials. One researcher thought that there was an issue concerning missing data; one noted publication bias towards positive results affecting systematic reviews; and another considered there was a need for more research into different ways of entering people into studies and more interim analysis of trial data.

One researcher was critical of certain biases inherent in the HTA Programme. It was suggested that there was too great an orientation towards randomised trials compared with other methods, and to quantitative compared with qualitative research. Another said that there was too great an emphasis on achieving certainty, arising (it was proposed) from the culture of medical professionals represented in the Programme:

> "The fundamental question of methodology is how to get an answer to the question that you are asking that is convincing to you and the other people who you want to convince … It has gone off on a more ideological tangent that says, 'This is what we need to know', without appreciating that this is really about belief rather than knowledge."

Several researchers thought that the Programme should turn its attention to issues of "implementation" or "development" (i.e. getting the results of research into practice):

> "It's how to bridge the gap between findings from randomised controlled trials … how to then translate appropriately the findings of that into appropriate clinical practice and policy making. Funding more and more systematic reviews or more and more randomised controlled trials doesn't necessarily improve healthcare and decision-making."

> "I keep saying to people: 'You can collect all that data, but really we wouldn't know what to do with it' … I'm worried about the HTA doing systematic reviews … and then no one getting round to putting it into practice. It would be very useful for them to fund workshops to discuss the methodologies' applications, sharing datasets … that would be quite a good model for other things."

One interviewee expressed the hope that the Programme would prepare a summary report concerning future research and circulate it widely among the methodological research community.

Another raised issues concerning the tendering process:

"There is a big problem about the whole commissioning process … where you spend ages putting in an outline bid and then it gets rejected or whatever. Although it gives everyone the opportunity to get funding, in the end it wastes a lot of people's time. I think it's fair to go down the line of a more limited tender for projects."

Chapter 4

The Methodology Programme: views of those involved in its creation and development

This section summarises the views of six people who played a key role in the development of the Methodology Programme, who responded to a letter with three principal questions.

Aims of the Programme and their achievement

The respondents were asked whether the Methodology Programme (Panel) had achieved the aims held for it at the time it was set up, as seen by themselves or others. For maximum clarity, they were also asked to indicate the nature of these aims as well as how the Programme had achieved them and what may have helped or hindered their achievement.

Views about the aims of the Programme did not differ widely, although two people questioned whether they had ever been explicitly set out. The aims were seen to be, as expressed by one man, "identifying and answering important methodology questions relevant to the practice of health technology assessment". If the programme were successful, he continued, "it would be easier for researchers to answer [health technology assessment] questions in ways that were both valid and relevant to decision-makers in the NHS." One person added that a further aim was to "identify gaps" in basic methodological research for the HTA Programme. The same person also suggested that the Programme was set up in part in recognition of the minuscule amount of funding provided to applied research for the NHS, compared with biomedical funding. Another thought that an aim had been to "raise the profile of health services research methodology".

One person, who was not involved at the outset, suggested that the clearest description of its aims was that on the cover of the BMJ book,[5] that is, to provide a practical guide to health service researchers to the approaches they should be taking. He argued that this was not a realistic aim, either for the book or for the Programme, because many researchers were only partially trained and would not have the technical background knowledge to cope with the reports. On the other hand, a *post hoc* rationale would be to inform those trained and engaged in high-level HSR about current trends in methodology. As such, it was directed to a very small clientele.

There were more mixed views about the success of the Programme. Four considered that it had clearly been successful. One wrote: "impressive in its depths, breadth and the extent of its coverage". The BMJ compendium and the planned Sage book[6] were seen as a testament to its success, as was the reception of many individual monographs. Another noted that he felt "immensely proud" that the NHS R&D Programme had been responsible for the Methodology Programme, described in the BMJ book as "the largest, most wide-ranging enterprise of its kind ever undertaken". He thought it was particularly noteworthy that it had attracted strong bids for work from researchers overseas. One respondent noted that the methodology reports produced to date comprised one-third of all those within the total HTA Programme, although there were six Panels contributing to its activity.

Among those who agreed that the Programme had been successful, there was considerable consensus that this stemmed from a combination of enthusiasm among the researchers, aided by the availability of funding for work they believed in, and steering from the top. One noted the Programme's ability to engage "top quality" health services and health technology assessment researchers, not only on the Panel itself but also in refereeing project proposals and reports, and undertaking the needed research. Another gave particular credit to the first chairman of the Panel: "His clear thinking, knowledge of the field and the research required, and enthusiasm enabled him to motivate a well-chosen panel." It was suggested that the Methodology Panel had been more successful than others (including a larger output to date) partly because of the greater preparation and knowledge of health services research among researchers in the field compared with those involved with other panels.

One person noted that quantitative research, particularly the design of clinical trials, had been

particularly strong, but, because there was extensive research globally, the Programme had had little scope to break new ground. Qualitative research, on the other hand, had been of a high standard, although under-represented, but philosophical diversity and fragmentation among qualitative researchers made it difficult to establish consensus on key methodological questions.

Two people referred to the role of the Methodology Working Group as a further stimulus to enthusiasm. One commented that, although the meetings of both the Methodology Panel and the Working Group appeared at times to be "a rather exclusive club", this very exclusivity had contributed to the Programme's success, because of the high level of enthusiasm for the work.

One respondent thought that the Programme had dealt effectively with three problems faced at the outset: (1) a lack of understanding within the NHS of what constitutes health services research and how it should be undertaken (together with scepticism about the research methods used); (2) a shortage of methods for addressing the research questions posed by the R&D Programme in general and the HTA Programme in particular (especially apparent in studies investigating interventional and diag-nostic techniques); and (3) a shortage of trained research workers in the field. This last view seems in direct contradiction to that noted above, that there was a good supply of researchers in the area.

However, some caveats were noted. One person thought that the Programme had been successful in prioritising methodological research needs, but less so in commissioning such work. He considered that there had been insufficient resourcing of research management, which was necessary because of a "culture change to needs-led research" rather than researcher-driven research. He also argued (confirming the view just noted) that there was a shortage of methodologists, exacerbated by the fact that some had been already hard-pressed with other work demands when the work commenced. Another noted that the BMJ book was not a very successful exercise, seen as an overpriced "advertisement for the monographs" because these were, in any case, free to researchers on the Internet.

One respondent added that the real test of the Programme would be the extent to which "the fund of knowledge that has been gathered and sifted" was used. This would require a different kind of exercise to ascertain.

Noteworthy projects

Secondly, the respondents were asked to identify those projects that had been particularly well executed or especially useful in developing an understanding of important issues.

Two sought to identify the characteristics of such projects, rather than individual studies themselves. Both tended to concentrate on key characteristics of the researchers involved, particularly the need for senior people who were knowledgeable and reflective.

One considered that the most well-received reports had one or more of the following features: (1) a multidisciplinary team; (2) a higher ratio of senior to junior research input than for other funded research; (3) the systematicity of review methods combined with the "insights of a 'think piece'"; and (4) a good understanding of the target audience (both researchers/methodologists and others interested in health technology assessment, including clinicians, managers and policy makers).

The other suggested that projects should: (1) be led by people who were well conversant with the field; (2) closely follow the commissioning brief; (3) avoid making "naive assumptions" about systematic reviews; (4) entail a scope appropriate to the topic ("sufficient breadth for the answers to be in context and sufficient depth to probe new areas"); and (5) be undertaken by people who understood the philosophy of the Programme. This person further suggested that projects that had not been well executed (including some not covered by the BMJ book) were those that aimed to cover too much, did not respond to the brief, were unaware of the needs of the HTA Programme, and presented polemic rather than analysis.

One person said his response to projects necessarily reflected his own interests, concerned with the reduction of bias in assessing the effects of health care. He found all the contributions in Part Two of the BMJ book to be interesting as well as chapters 14 and 16. He also noted that the two appendices contained a number of mistakes, which could have been easily rectified by wider consultation prior to publication.

Two respondents suggested that those projects that had dealt with the area of randomised trials were particularly valuable. One argued that the relevant chapters should be "compulsory reading" because they set out the evidence for and against randomised controlled trials over the full range of issues

confronting those undertaking them. The other cited virtually all the chapters in Part Two of the BMJ book. He was particularly impressed with the project on consensus development methods.

Disappointment was expressed by one person concerning the chapter on a Bayesian approach because it left the reader "interested but incapable of applying the contents to personal experience". This was particularly problematic given current encouragement to adopt such an approach, coupled with a high degree of scepticism among those who have not used it. The publication of the HTA Series monograph of this research became particularly urgent in the light of these considerations and he hoped that the Methodology Programme could expedite the process. A comparison was made with the introduction of multiple least squares regression in the 1960s and logistic regression in the 1970s, both of which took time to be assimilated by the research community.

One respondent added that "the whole is greater than the sum of the individual parts", suggesting that the book would add to the "respectability" of health services research and that this would be further enhanced when the methodology manual (i.e. the planned Sage book) was produced.

Future directions for the Methodology Programme

Finally, the respondents were asked how, in the light of its achievements to date, they would like to see the Methodology Programme develop in future. All the answers were quite different and need to be set out in turn, although there was frequent emphasis on the need to address the dissemination and use of research.

One person questioned whether the Programme should continue to commission methodological research in the field of health technology assessment. He noted a striking change between 1993 and 1998 in the nature of topics considered for commissioning. The latter were highly complex, to the extent that probably only half of the Panel members could really understand them, whereas, in the early days, all members understood all topics. He posited that there may be diminishing returns in investment in health technology assessment methodology questions, at least for a period. Instead, the Programme should place its energies in disseminating the results of its own (and other) methodology research to health technology assessment researchers, followed by "a lengthy listening exercise". He stressed that he was not proposing a moratorium on all methodological research, but only health technology assessment research, to concentrate limited research funds into the Service Delivery and Organisation programme and "other more unploughed territories".

Another expressed concern about the potential effects of the change in portfolio, with new studies commissioned to support the Service Delivery and Organisation and the New and Emerging Applications of Technology programmes, as well as demands from other arms of the NHS. The success of the Programme had derived from the structure of the Panel and relationships between the chairman, panel members and those commissioned to do the work, but the system could become "less harmonious and thus less efficient" with a wider portfolio. It was suggested that the new Panel sitting within the new independent Methodology directorate should be evaluated to see if it still works as well.

One respondent was particularly concerned with aspects of practical research management. He suggested that bids should provide evidence of senior staff involvement in topics, as well as cover a range of methods. Moreover, there should be greater hands-on management on the part of commissioners. He proposed that researchers should be invited to submit a draft report for early consideration "before the work is too far under way", enabling an expert committee to insist on some redirection where necessary. Furthermore, there should be an obligation on researchers to state their conclusions with a clear protocol and summary of the work undertaken. The latter should be no more than five pages, to enable the broad conclusions to be readily discussed with others in the field.

A few substantive suggestions about what the Programme should commission were proffered. One person noted that the principal need was for systems for updating reviews as new evidence becomes available. He also raised the problem of systematicity in the context of methodological work, and hoped that substantial numbers of social scientists and statisticians would in future collaborate in such projects.

Another suggested a focus on two themes. First, there should be attention to how to make better research use of the vast NHS activity; he proposed that it should be possible to derive generalisable information at lower marginal cost and with greater external validity than the conventional research designs currently used. Secondly, specific

methodological gaps in the Service Delivery and Organisation agenda should be identified. He questioned why this lacks a stronger research base and what could be done about it, given the favourable NHS environment.

Yet another respondent considered that, having set the scene with its earlier research, the Programme should create a list of methodological problem areas and commission more primary research. Two possible areas were proposed: first, developing computer software and exploring specific practical problems relevant to the application of a Bayesian approach, and, secondly, the development of sound approaches to reliable and valid measures of patient (or other) opinion. In addition, there was a need to tackle problems arising in earlier research, including the characteristics of failure – whether general or local – to recruit patients into studies.

One person suggested that the Programme would need to question two assumptions that are inherent in the culture of health services research. These were: (1) that research could be undertaken by junior researchers, with some senior management; and (2) that all reviews should be seen as classic Cochrane systematic reviews (felt to be more an assumption of the researchers themselves than of the Programme itself). In contrast, he considered that such reviews should be seen as "part data gathering exercises, part selective reviews and part think pieces".

Another argued that, despite the recent decision to broaden the role of the Methodology Programme beyond health technology assessment, "there remains a continuing problem in defining what the content of the work should be and who the customers for its outputs are." He thought that methodology had no "easily defined constituency separate from the totality of applied health research", so that its outputs were an incremental addition to the broad collection of knowledge, rather than being of distinctive interest to the NHS.

Chapter 5

Interest in the methodology projects

This section summarises the data from the analysis of use of the NCCHTA website, using data from the end of February and the end of May 1999 as well as use since the publication of individual reports.

Website analysis

In all, 29 projects were published, at a fairly steady rate, between February 1997 and the end of May 1999, ten of which focused on methodology and 19 on other issues. The former were published slightly later (seven published after October 1998) than the latter (14 published before October 1998).

By the end of February 1999, the median number of hits was approximately 1200 for methodology projects and 800 for other projects; both medians had increased by 300 by May (*Table 1*). Although hit count alone reflects interest and, to some extent, length of time since publication, the hit rate per month since publication allows an adjustment for the different dates of publication (*Figure 1*). The median hit rate for the seven methodology projects (155/month) was more than double that for the 17 other projects (66/month) by the end of February 1999 and nearly triple (241/month versus 89/month) by the end of May 1999. These differences are significant at a 5% level, using the Mann–Whitney test.

The number of hits per month up to the end of May 1999 increased with calendar time, so the most recent publications had the highest monthly hit rates (*Figure 1*). However, the rate of increase differs according to type of project. For methodology projects, the increase is 12.5 hits per month, whereas, for other projects, the increase is 4.5 hits per month.

If one considers only the 24 projects published before March 1999, the contrast in the increase in hit rates per month to May 1999 is greater: 23 for methodology projects and five for other projects. This relationship is stronger with log hit rates per

TABLE 1 *Summary statistics for hits on the HTA website*

	No. projects	Minimum	Lower quartile	Median	Upper quartile	Maximum
No. hits by February 1999						
Methodology	7	309	587	1159	1494	1506
Other projects	17	91	497	783	1175	1873
Combined	24	91	514	912	1220	1673
No. hits/month by February 1999 (HPM Feb)						
Methodology	7	33	77	155	374	753
Other projects	17	18	38	66	105	329
Combined	24	18	41	80	163	753
No. hits by May 1999						
Methodology	10	49	1004	1545	2042	3342
Other projects	19	65	619	1108	1593	2556
Combined	29	49	722	1106	1677	3342
Hits/month by May 1999 (HPM May)						
Methodology	10	49	84	241	385	689
Other projects	19	23	47	89	116	234
Combined	29	23	55	100	192	689
Hits/month Feb – May 1999 (HPM Feb – May)						
Methodology	10	16	157	264	374	612
Other projects	19	22	59	102	159	294
Combined	29	16	65	136	264	612

month, indicating a 2.8% increase in hit rate per month; this means that, on average, the second of two projects published 1 year apart has a 40% higher hit rate than the first (*Figure 2*). The increase in log hit rate with month of publication is the same for both types of project, but the methodology projects have a higher initial level.

The difference between the hit rates for publication in March to May 1999 and the hit rates from publication prior to the end of February 1999 do not generally change with date of publication (*Figure 3*). There are five projects (two methodology and three other) with a substantial drop of at least 100 hits per month. The remaining 19 projects had an average increase of 90 hits per month.

From these data, some points can be noted. First, the rate of hits for the methodology projects is high in itself, with half the projects experiencing between 80 and 390 hits per month. Secondly, this rate is considerably higher than that of other projects within the HTA Programme. This can be assumed to reflect greater interest in the topics commissioned with a methodology brief, but one must then question why this may be so.

The Methodology Programme was particularly innovative and it may be that this in itself raised curiosity about its outputs. To the extent that this is correct, it suggests that such outputs were well publicised, possibly by word-of-mouth through a network of researchers and others interested in methodological analysis. In addition, however, the audiences for the two types of projects may differ. Some predefined groups among those likely to be interested in the "other" category might have been sent hard copy reports automatically, so they had less need to access information via the HTA website. In contrast, those interested in methodology may be less clear *a priori* and are therefore required to rely more on initiative and publicity to find out information. Finally, it may be that methodologists have greater familiarity with – and access to – the Internet.

Whatever the explanation, it can be cautiously proposed that the absolute level of interest in the methodology projects is clearly high, with the hit rate increasing from February to May 1999 for most. Regrettably, data are not available on the views of readers of the reports and other information obtained from this source.

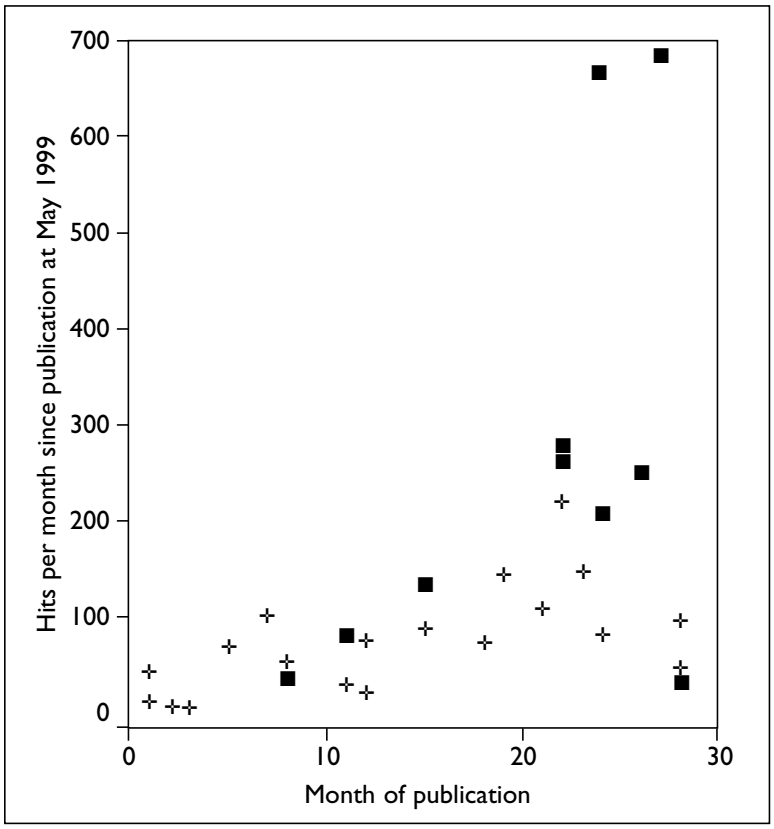


FIGURE 1 *Hits per month since publication at May 1999 against month of publication (■, methodology; ✛, other)*

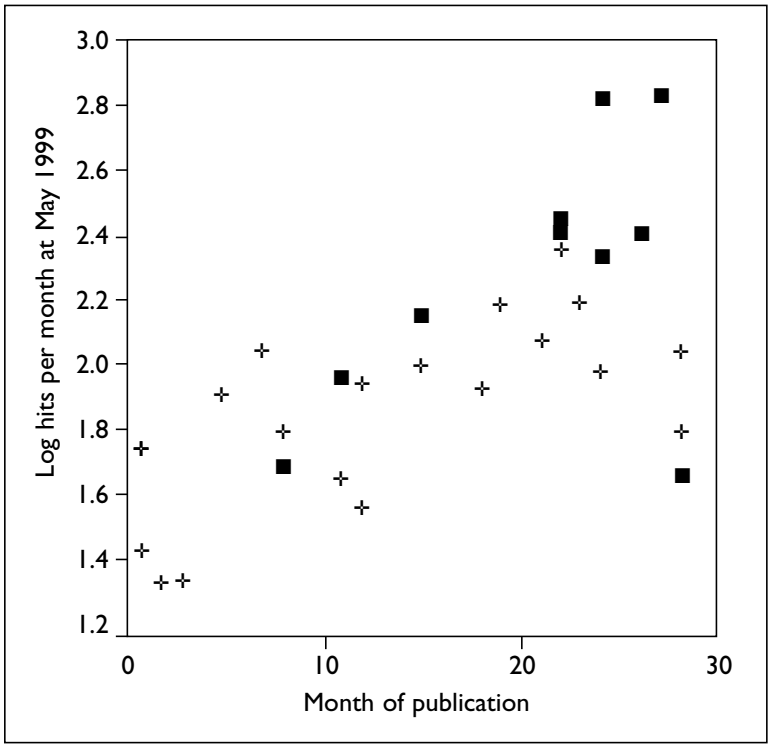


FIGURE 2 *Log hits per month since publication at May 1999 against month of publication (■, methodology; ✛, other)*

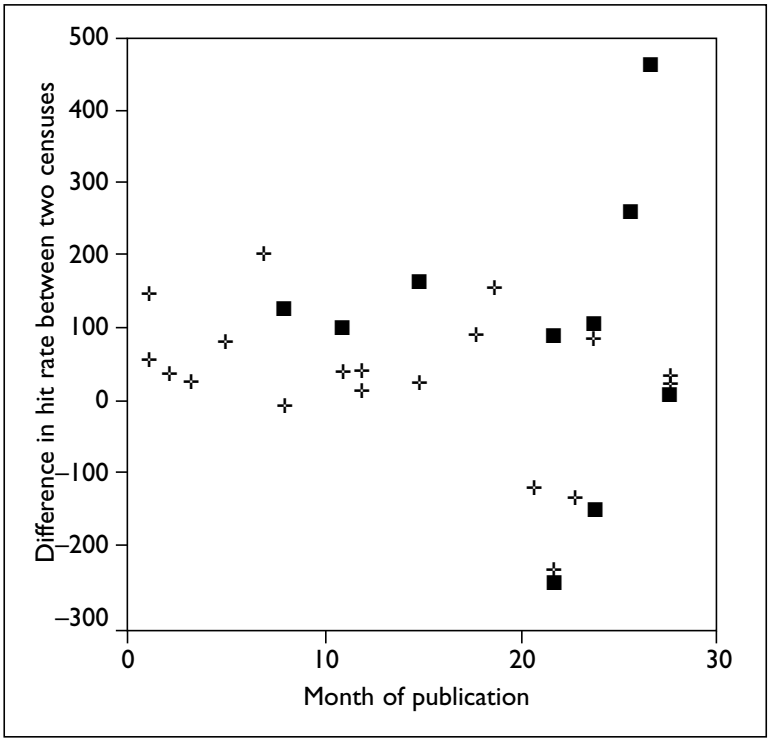


FIGURE 3 *Changes in hit rate according to month of publication (■, methodology; ✛, other)*

Chapter 6

Conclusions and recommendations

Reflections on the findings

One historical feature of the Methodology Programme bears heavily on the results of this study. The reviews on which it was based were commissioned at the zenith of enthusiasm for systematic reviews and meta-analysis. This seemed to flow over into the Methodology Programme, with the requisitioned reviews all intended to be "systematic". As has been shown, some researchers felt compelled to try to follow standards for systematic reviews of comparative quantitative studies, such as those in the Cochrane "tool kit",[4] including the need for an exhaustive search of all relevant literature. Yet, when applied to some topics, such as the ethics of clinical trials, this could lead to a deluge of repetitive material.

Many researchers concluded that the word "systematic" in the context of a methodological review should be interpreted as an organised, protocol-directed product, rather than an exhaustive search for every last paper on the topic. The concept of "theoretical saturation", whereby further searching of a certain type of literature fails to yield any new insights, has now gained broad currency. Our findings confirm the need for systematicity rather than exhaustiveness.

We had hypothesised that most studies would synthesise argument from the literature but that some, like systematic reviews of treatment regimens, would use the literature as data on which further analysis could be based, and that others might use other forms of data collection (such as surveys) or pure reason. There did not, however, appear to be much need for additional research methods in the research undertaken; there was little emphasis on triangulation. We found that all the projects studied used the literature, but only two used it as data on which further primary research could be based. No project concentrated on other forms of data collection (five noted some additional research methods) and none claimed to be based on pure reason.

However, hints at other methods have arisen: (1) from later projects, one of which was used to pilot this study; (2) respondents who built up a theory to "sort" the literature; (3) a study that had its argument built around a case study; and (4) five reports noting the use of additional research methods. On the whole, although researchers were not opposed to the idea of changing methods in the light of thinking afresh about a subject (in contrast to changes in research topics, which were seen as inappropriate), this was not a major issue for them. We are aware that projects commissioned in the years subsequent to those studied did make greater use of triangulation.

In retrospect, it seems likely that any finding of such diverse approaches was constrained by the particular set of projects studied. These were all commissioned in 1993 when the paradigm was of a "pure" review and when it was considered that many questions needed a scoping review. In subsequent years, some HTA methodology commissions have required new data collection.

Many other stimulating issues were raised by our review of methodological topics. For example, the question of whether such reviews should be seen primarily as data collection or as thinking exercises pervaded the interviews and has important practical implications. To the extent that they should be viewed as opportunities to generate some thinking by senior researchers, there is a need to build in sufficient senior staff time (which must be notably higher in these studies than in most primary research).

The ever-present issue of researcher objectivity also raised some interesting comments. We note that most researchers believed that some bias was inevitable, particularly in non-quantitative studies of this kind, but they nonetheless made careful efforts to reduce its impact. We were impressed by the variety of methods adopted for this purpose.

In practical terms, the organisation and management of the projects studied seemed quantitatively but not qualitatively different from those of other research. The need to plan out a research project carefully, including building in time for analysis and writing, and for senior researchers to keep a close eye on research management, while common to many studies, was heightened by the particularly vexing nature of the search and synthesis process in methodological reviews.

The evidence is, overall, that the Methodology Programme has proved to be a success. It has covered a wide range of issues and generated an impressive degree of enthusiasm among researchers. Whatever the problems along the way, it has managed to overcome a lack of understanding within the NHS about methodological issues and helped to build up a supply of researchers familiar with the field. Furthermore, our survey of hit rates on the HTA website demonstrates the degree of interest in the methodology reports.

Finally, it is our view that the significance of the Methodology Programme should be seen to lie not simply in the reports produced but also in the diffusion of knowledge it has facilitated. Many ideas were clearly stimulated and passed on via new and existing research networks (including discussions between researchers on projects with overlapping briefs) and other meetings (such as those of the Methodology Group) that were set up directly or indirectly as a result of the Programme's existence.

Reflections on the methods

This research entailed three principal research methods: telephone interviews, correspondence with key people involved in the development of the Methodology Programme, and an examination of the use of the HTA website. The soundness of these approaches could reasonably be questioned and deserves brief attention.

We believe that the telephone interviews worked very well; they elicited very full responses and much valuable material. Those interviewed seemed to find them stimulating and we consider that we obtained the necessary information. The fact that they were tape recorded and transcribed meant that they could be analysed in detail. It can be added, however, that they were time-consuming and tiring for both researcher and respondents. Questions were sometimes long and convoluted and, on more than one occasion, needed to be repeated (and sometimes rephrased) before the respondents felt able to answer. Nonetheless, we would have no hesitation in suggesting that such a method was worth while for this kind of study.

The letters to key people were clearly less successful. There was a very poor response rate. This might have been expected from very busy people, but it could have been exacerbated by the fact that the letters competed with other demands on potential respondents prior to summer holidays. In addition, the request that those involved should read a fairly lengthy compendium prior to replying added to its onerous nature.

One person also questioned the method in terms of achieving the desired result:

> "Relying on written responses to open questions seems on the face of it likely to be a method of obtaining views that lacks richness, will discourage people from giving full responses and may inhibit dialogue and understanding."

Whether this is correct or not was not well tested because of the paucity of the response, but most replies were reasonably full and showed that considerable thought had been given to the issues. Indeed, it might have been the difficulty of replying quickly that was most discouraging to those approached.

In terms of the website analysis, we would note that it is important not to make inferences regarding either the quality or the impact of research on the basis of hit rates, in much the same way as it would be problematic to make inferences about other research from citation analysis.

Recommendations

A number of recommendations suggest themselves as a result of this study. Some of these pertain directly to researchers (although with implications for the Methodology Programme as commissioner) and some concern the future planning of the Programme, with respect to both the management of researchers and the content of future projects. These are set out, in turn, below.

Recommendations for researchers

The conduct of research

Reviews on methodological topics are very different to typical "Cochrane type" reviews because they are effectively a form of qualitative research, with all the complexities of analysis that this entails. Some implications are:

- Investigators should not aim to chase every last reference, but should ensure that they search widely in disparate databases and sources.
- Researchers could consider methods beyond the review of ideas and even the review of data, for instance: intellectual analysis, networking and primary data collection such as in methodological studies attached to primary health technology assessment, and other methodological experiments.

- Clear separation of the various stages of research, including data collection, analysis and writing, may often prove unhelpful and a greater degree of overlap than is usual should be accommodated in the work programme. In particular, commencing analysis early can help to clarify data collection needs and commencing writing can clarify analysis and synthesis.
- Researchers should publicise the existence of their studies as broadly as possible; this will bring in new ideas and possibly short-circuit extensive search processes.
- All studies should include a short summary of key findings, which should include practical solutions to identified problems, to assist future researchers.

Reducing bias

The potential for bias is critical in methodological research because many arguments depend on underlying philosophical assumptions or other perspectives. A variety of safeguards should be built in. Most of these are generally used by projects; they include:

- the establishment of a steering group, to ensure that premises are widely shared and that the arguments are coherent
- the involvement of multidisciplinary teams
- peer reviewing of both project applications and final reports
- the preparation of a report for the original commissioner (the Methodology Programme in this case), so that decisions can be taken on whether there is a wish for general "ownership" of the results and any policy implications can be considered
- a clear intention to publish in widely disseminated journals, in order to attract further comment.

The efficient management of research

Some of the above recommendations, for instance, setting up a steering group, may also help to ensure more efficient research management. In addition, we recommend:

- Senior staff need to be involved throughout these research projects. This is partly because much methodological work is analytical, rather than simple data collection, and partly because good research management is essential. This should be reflected in the costing of grants.
- A clear plan of action and research timetable should be developed, even if it is changed in the light of new thinking during the course of research. Planning the report writing is particularly important, not only for deciding who should do this and when but also for devising a preliminary structure to influence both data collection and analysis.
- Projects with many sites experience difficulties, so the dynamics of such projects should be examined carefully prior to their funding. This should include both the logistics and the willingness of the parties to work together.

Recommendations for the Methodology Programme

The Methodology Programme may be seen to be at a crossroads, having commissioned a substantial number of reviews and used this research as a means of assessing future directions. Some suggestions for the management of the Programme include:

- The Programme should continue the style of research management that it has used to date. This entails a fairly light involvement, coupled with an accessibility to researchers if there is need to discuss problems or proposed changes to the initial plan. Programme representation on a project steering group is one means of achieving this.
- It should also continue to assist projects to work together as needed, particularly where their briefs are similar or overlap in some way. They should be fully informed about other projects currently being funded and some resources toward the costs of meetings should continue to be provided. A master steering group for all projects would be one way of achieving this and reassessing resource needs.
- Prior to commissioning projects, the Programme could introduce a more iterative approach to ensure that it has correctly specified the research problem.
- The dissemination and use of the findings of the methodological research already commissioned by the Programme (and commissioned through other sources) should be given particular attention.
- The need for mutual support of those researchers undertaking complex projects should not be overlooked; the Methodology Projects Group played an important role in this respect and funding should be provided for this.
- The proposed wider Methodology Programme could be evaluated to see if it continues to work as well as it did when its brief was more limited.

We are conscious of the need to expand the types of methodological research commissioned. Such

research may range from intellectual analysis, through reviews of what others have said, to primary research, which may itself take many forms. Methodological research can be undertaken in conjunction with substantive research, for instance, comparing the results from different methodological approaches to a single topic. It can also encompass pure methodological experiments, such as eliciting patients' understanding of the rationale for studies in which they are asked to participate. Some suggestions about what could be commissioned include:

- Methodological researchers should be encouraged to explore the potential of using a wider range of methods, as outlined above.
- Some substantive researchers should be encouraged to add a methodological component to their studies, but not all subject areas can work within such a framework.
- The management of methodological research could itself be the subject of study.
- Arrangements could be set up for updating reviews as needed in the light of new evidence; those undertaking the initial review should be approached first for this task.
- Particular attention could be given to methodological gaps in the Service Delivery and Organisation agenda.

Acknowledgements

The authors wish to thank the contributors to the NHS R&D HTA Programme whom we interviewed for their patience and their thoughtful responses. We also thank those involved in the design and running of the Programme who were able to respond to our questionnaire.

References

1. Tannenbaum S. Knowing and acting in medical practice: outcomes research. *J Health Polit Policy Law* 1994;**19**:27–44.
2. Harrison S. The politics of evidence-based medicine in the United Kingdom. *Policy Polit* 1998;**26**:15–31.
3. Ferlie E, Wood M, Fitzgerald L. Some limits to evidence-based medicine: a case study from elective orthopaedics. *Quality Health Care* 1999;**8**:99–107.
4. Irving M. Foreword. In: Black N, Brazier J, Fitzpatrick R, Reeves B, editors. Health services research methods: a guide to best practice. London: BMJ Books; 1998. p. xiii–xiv.
5. Black N, Brazier J, Fitzpatrick R, Reeves B, editors. Health services research methods: a guide to best practice. London: BMJ Books; 1998.
6. Stevens AJ, Abrams K, Brazier J, Fitzpatrick R, Lilford R, editors. The advanced handbook of methods in evidence based healthcare. London: Sage; 2001.

Appendix 1

The projects studied

Assessing the quality of randomised controlled trials: implications for the conduct of meta-analyses
Mr David Moher, Children's Hospital of Eastern Ontario

Assessing and increasing the generalisability of randomised controlled trials
Dr David Henry, University of Newcastle

Factors that limit the quality, number and progress of randomised controlled trials
Dr Robin Prescott, Edinburgh University

Systematic review of literature comparing the use of randomised controlled trials and quasi-experimental or observational methods for assessing effectiveness of interventions and quality of care
Dr Barnaby Reeves, Royal College of Surgeons

A systematic review of randomised controlled trials and observational methods
Professor Martin McKee, London School of Hygiene and Tropical Medicine, University of London

Systematic reviews of trials and other studies
Professor David Jones, University of Leicester

Quality of life assessment and survival data
Dr Keith Abrams, University of Leicester

A purchaser's guide, review and archive of literature on qualitative methods
Dr Elizabeth Murphy, University of Nottingham

A review of the use of health status measures in economic evaluation
Mr John Brazier, University of Sheffield

Evaluating approaches to the costs of health technologies
Professor Martin Buxton, Brunel University

Handling uncertainty and variability in cost data when performing economic evaluations
Dr Alastair M Gray, Oxford University (no interview)

Development of explicit criteria for evaluating patient-based outcome measures
Professor Ray Fitzpatrick, Oxford University

Design and using patient and staff questionnaires: a review of best practice
Ms Elaine McColl, University of Newcastle

Statistical methods for health technology assessment: establishing good practice and identifying opportunities for innovation
Professor Deborah Ashby, St Bartholomew's and the Royal London School of Medicine and Dentistry

Bayesian methods in health technology assessment
Dr David Spiegelhalter, Cambridge University

When and how to assess fast changing technologies: a comparative study of medical applications of four generic technologies
Dr Jane Bower, University of Aberdeen

A systematic review of consensus development methods
Professor Nicholas Black, London School of Hygiene and Tropical Medicine, University of London

Factors that influence the effects of health technology assessments in practice
Mr Nick Freemantle, York University

Ethical issues in the design and conduct of randomised trials leading to an ethical framework for health technology assessment in the NHS
Professor Richard Lilford, University of Birmingham

Implications of socio-cultural contexts for ethics of clinical trials
Dr Jane Hutton, University of Newcastle

Appendix 2

Telephone interview schedule for researchers

Thank you for agreeing to answer these questions. As you know, most of them concern your project [*Ref. no. and title*] with a few more general questions.

Please would you try to be as thorough, accurate and honest as possible. Some of the questions I will be asking are quite specific, but the interviews are not going to be linked to you, your research or your colleagues in the final report.

I will talk quite quickly as we have a lot to fit in, but do tell me if you would like me to slow down. I am now recording this conversation. Do you consent to this?

We will start with a couple of quite **General** questions.

1. Do you mind if I ask you, first of all, what stage your research is at just now?
   a) Research continuing
   b) Research completed, writing-up
   c) Completed, with referees
   d) Changes being made following referees comments
   e) Awaiting publication
   f) Published

2. How many people were actively and extensively involved in your project?

3. What research disciplines are you all from? By that I mean what research discipline did each of you bring to this project?

Now we'll move on to the **Research Design and Objectives** section.

4. This may seem like a difficult question to start with, but what type of study do you see this research as?
   a) Systematic review
   b) Some other type of review
   c) Not a review, doing something else [*Probe deeply!!*]

The next section is entitled **Research Topic**. Later we will turn to your methodology, so it would be helpful if you could think about your research topic and your research method quite separately.

5. We could say that most researchers within the Methodology Programme found that they **had** to move their project away from their brief. Did you have to change **your** topic at all during your study?
   a) Yes
   b) No [*If 'no', go to question 8*]

6. Some researchers changed topic completely by moving from their original topic, others changed less dramatically by becoming more focused on a specific area; some even changed topic by extending their original focus. In what way did you change your topic?

Alternative 6.

So you didn't change your topic as a whole. I wonder if you had to change anything about your topic? It seems as if several projects had to become more focused on specific areas of the original topic or expand the focus to look at more than the original topic.
   a) Narrowing the topic by investigating only one angle of it [*Precision*]
   b) Reducing the depth of the topic by offering an overview of each area
   c) Moving across from the original topic [*Lack of accuracy*]
   d) Extending the original topic

7. Why did you change your topic?

Now we're going to move from what you studied to how you studied it, in this, the **Methodology** section. I would like you to think about methods in the broadest sense, so that we talk about everything you did.

8. What did you expect the outcome of your study to be at the outset?

9. Was your prediction correct?

10. What was the main kind of evidence your review used?
    a) Data collected through a literature search
    b) Data collected some other way
    c) Even split of methods

[*NB If they did not use a literature search go to question 17.*]

11. We can distinguish between literature searches that are exhaustive and those that are systematic but not exhaustive. Was your literature search exhaustive, systematic but not exhaustive, or is it better described some other way (e.g. selective)?
    a) Exhaustive
    b) Systematic, but not exhaustive
    c) Something else (e.g. selective)

12. Did you design and use a preplanned literature search strategy to help you to limit or maximise your search?
    a) Yes – preplanned search algorithm
    b) No – algorithm designed by trial and error – why?
    c) No – recognised the algorithm only in retrospect – why?
    d) Other – specifics

13. How did you identify sources?

14. Did your literature search generate too many or too few papers in the first instance? (That is, was your initial search too sensitive, too specific, both, not sensitive enough, not specific enough, both, just right?)

15. [*If found too many or too few relevant papers*] How did you cope with that?

16. To what extent did you acquire important references by chance?

17. Did you use other (additional/alternative) method(s) to collect data other than literature?
    a) Yes
    b) No

18. How did you arrive at them/it?

19. It seems that many researchers in the Programme had to introduce methods that they had not envisaged in the protocol. Were your additional methods preplanned? [*Specify which one(s)*]
    a) Yes [*If 'yes', go to question 21*]
    b) No

20. Why did you change your methodology?

21. Did you attempt to weight studies in your synthesis?
    a) Yes [*If so, how: by study design, sample size, other*]
    b) No [*If not: why was that?*]

22. How would you describe the methods you used to analyse or synthesise the literature (or data) you had gathered?

23. Was the process of analysis and synthesis as you had expected?

24. Were you able to draw concrete conclusions as a result of your study?
    a) Yes – how?
    b) No – why not?

25. Were you able to make specific recommendations in the light of what your conclusions told you?

We now turn to **Project Management**. I will particularly ask you to make a few comments with hindsight here.

26. Were there any problems with allocating time and effort among the following?
    a) Planning
    b) Data collection
    c) Analysis
    d) Write-up
    e) Other activities

27. Was there anything unusual or unexpected about the way you had to divide up your time between these activities?

28. Did your project take longer to complete than you had stated in the protocol?
    a) Yes
    b) No

29. Why was that?

30. How would you describe the approach that was adopted to managing the day-to-day running of the project? (What were the main strategies in terms of supervision and communication?)

31. Did you learn lessons from this project about management or feasibility? (e.g. in connection with division of tasks between researchers, scheduling of work, allocation of resources and time)

32. What general advice would you give to researchers or research groups doing the same kind of study in the future?

33. Are there any additional methods of research or data collection that you would now use, with hindsight?

34. Are there any additional methods that you would now use, with hindsight, for synthesis or analysis?

Now it would be good to look at **Who Became Involved** with your research. We will start with the role of the Department of Health, R&D or the NCCHTA.

35. If/when you (had) found that you could not keep to your stated research protocol (contract), would some formalised external steering have been helpful to you?
    a) Yes
    b) No [*If 'no', go to question 38*]
    c) Maybe

36. Why would that have been helpful?

37. Who would it have been helpful from?
    a) Steering committees (composed of …)
    b) Other (specifics)

38. Do you think the Programme exercised an appropriate level of input to your project in general? Would you have welcomed more/less intervention and help?
    a) More
    b) Less

39. Were referees' comments helpful?

40. How many referees would have been a good number for this kind of project?

41. Did the referees have the right skills?

42. What skills should the referees have had in your case?

Now we turn to **Other Contacts** outside your project group.

43. Did your research overlap with work by any other HTA commissioned group?
    a) Yes
    b) No

44. Did you collaborate with any other groups within the HTA Programme?
    a) Yes [*If 'yes', state who, how and how much*]
    b) No [*If 'no' go to question 46*]

45. How helpful or important was that to the project?

46. Would it have been helpful or unhelpful to you if collaboration had been more actively coordinated?

47. Did you go to meetings of the Methodology Group regularly?
    a) Yes
    b) No

48. Was that helpful?

49. Did you collaborate with any groups outside the HTA Programme?
    a) Yes [*If 'yes', state who and how/how much*]
    b) No [*If 'no' go to question 51*]

50. How helpful or important was that to the project?

This section asks about **Who was Directly Involved** in your research and whether the allocation of funding was beneficial to your project.

51. Which of your original applicants actually played a part in the study?

52. Would your project have benefited if more senior staff had been more actively involved? (e.g. for employment of more senior staff, using substitution funding and more costing for consultancy time)
    a) Yes – How senior? How much involvement?
    b) No

53. Did your steering committee or decision-making group extend beyond the grantholders? Was your project guided by anyone external to the project grantholders?

This section explores **Bias, Subjectivity and Changes to Objectives**. First, a very general question:

54. Do you think it is possible to achieve a value-free presentation of quantitative or qualitative results? (i.e. neutral, objective)
    a) Yes
    b) No

55. Did you do anything to make your results less value-laden?

56. Do you think the peer-review process was helpful in making your review less value laden?

57. Do you think the peer-review process was helpful in identifying gaps?

58. In general, do you see changing the research topic once a project is under way as a problem:
    i) For the HTA Programme/process?
       a) Yes
       b) No

ii) For the validity of the study?
c) Yes
d) No

59. In general do you see changing the research method once a project is under way as a problem?
i) For the HTA Programme/process?
a) Yes
b) No
ii) For the validity of the study?
c) Yes
d) No

Now we need to take a look at **The Future of Your Study**

60. What has been done with your findings? Where has the work or parts of it been published? [*If they cannot remember, offer to e-mail about this.*]

61. I would like you to think about what should be done with your results. Apart from the normal methods of dissemination, such as monographs and journal articles, are there any specific ways that your results should be used? For example:
a) Fed into professional training. Training for whom? How?
b) Added to checklists for research councils.
c) Particular types of dissemination (e.g. to purchasers). Dissemination to whom? How?
d) Change NHS practice. What changes? How?

62. Who do you see as the primary audience of your study?

63. Did you identify any ongoing or further research issues?
a) Yes
b) No

64. Do you think any new research should be funded as a priority as a result of your study?
a) Yes
b) No

I have a few questions that ask about your opinions on **Commissioning**.

65. On balance, is more progress going to be made in HTA methodology now by bolting methodology research on to other substantive studies or by maintaining methodology research in its own right? (i.e. continuing to seek answers to methodological questions as the sole purpose of pieces of research versus incorporating methodology research into studies that are already ongoing)
a) Bolt-ons
b) Dedicated studies
c) Combination
d) Other

66. I have a similar question about small studies, by which I mean 6-month scoping projects, and large studies, by which I mean extended 3-year exercises. On balance, is more progress going to be made in HTA methodology now by funding more small studies or more large studies?
a) Small (scoping exercises of less than 6 months)
b) Large (roughly 3 years)
c) Combination
d) Other

67. Again on a general note, do you think reviews of methodology subjects should be updated at some stage after initial publication?
a) Yes
b) No [*If 'no' go to question 69*]

68. How should this be done?

69. Has the HTA Methodology Programme covered the big issues?
a) Yes. What were they?
b) No. How should it go about identifying the important questions? or What are they?

70. In general what do you think is the best method or combination of methods for methodology research?
[*If they are struggling with this ask if they see a systematic review as the most appropriate method.*]

Ask principal investigators only:
We are hoping that one of the more tangible outcomes of our project will be a methodology database. Therefore would you and your colleagues be happy for the database of references that you put together for this project to be made available to other researchers as part of that?

[*Get Research Assistants' addresses where applicable.*]

Appendix 3

People approached for their views on methodological research

Professor Martin Buxton
Professor of Economics
Brunel University

Professor Iain Chalmers
Director
UK Cochrane Centre, Oxford

Professor Antony Culyer
Deputy Vice-Chancellor
University of York

Professor Charles Florey
Head of Epidemiology
University of Dundee

Professor Chris Henshall
Deputy Director
NHS Research & Development
Department of Health, London

Sir Miles Irving
Chairman
Newcastle upon Tyne Hospitals NHS Trust

Dr Ruairdh Milne
The Wessex Institute for Health Research
& Development
University of Southampton

Professor Sir Michael Peckham
University College London

Professor Ian Russell
Department of Health, Sciences and
Clinical Evaluation
University of York

Professor Trevor Sheldon
Director
NHS Centre for Reviews and Dissemination
University of York

Professor Andrew Stevens
Department of Public Health & Epidemiology
University of Birmingham

Professor Kent Woods
Director
NHS HTA Programme
Professor of Therapeutics
University of Leicester

Appendix 4

Letter to people active in the Methodology Programme

Dear …

Study of Methodology Research

I am writing to ask your assistance with a research project commissioned by the HTA Programme. I am the lead grantholder, but the grant is held jointly with Professor Raymond Fitzpatrick, Professor Andrew Stevens, Dr Jane Hutton and Dr Sarah Edwards. It was set up to review some of the epistemological and management difficulties experienced by HTA-funded methodology researchers, to consider the value of their research to potential users and to explore some ideas about the future of the Methodology Programme.

We have now completed the first of these tasks, involving telephone interviews with 35 researchers, and have prepared a draft report on their responses. But instead of undertaking a series of interviews with potential or actual users of their research, we have concluded that it would be more helpful to obtain the views of key figures in the field on the success of the Programme to date and how it should be taken forward.

I am therefore writing to ask if you would be so kind as to write your thoughts on the following questions:

i) Has the Methodology Programme achieved the aims you held for it at the time it was set up (or which others may have held)? What were these aims, how has it achieved them and what do you think helped or hindered their achievement?

ii) Do you have any views about which projects were particularly well executed and/or particularly useful to developing our understanding of important issues?

iii) In the light of its achievements to date, how would you like to see the Methodology Programme develop in future? For instance, should it be expanded or contracted and what kinds of research should it concentrate on?

Your answers will be treated as research data, i.e. anonymously, so that you should feel able to express your views freely. Indeed, the information will be analysed by Dr Ann Richardson, a freelance senior researcher who has been helping us with this research; as in the case of the interview data, your responses will not be passed to us except in an anonymous fashion.

In order to assist you to consider the Programme as a whole, you might like to look over the BMJ-sponsored book *Health Services Research Methods: A Guide to Best Practice*, edited by N. Black *et al.,* BMJ, 1998, which provides a useful summary of the output of the methodology work to date. If you do not have a copy, please do let me know and I will send one to you very quickly. We would certainly prefer that your comments were drawn in part from a perusal of its contents.

We would be very grateful if you could respond to these questions by post or e-mail to Ann Richardson, details given below, within the month, say by Friday 2nd July 1999. You should not feel compelled to write long essays, but of course we hope you will write enough to convey your thoughts as clearly as possible.

If you have any prior queries, please phone me or Ann Richardson and we will endeavour to help.

Yours sincerely,

Professor R J Lilford, PhD, FRCOG, FRCP, MFPHM
DIRECTOR OF RESEARCH & DEVELOPMENT
NHS Executive, WM Regional Office
PROFESSOR OF HEALTH SERVICES RESEARCH
The University of Birmingham
NHS CLINICAL TRIALS ADVISOR
DIRECTOR OF CENTRALLY COMMISSIONED
PROGRAMME ON METHODOLOGY RESEARCH

replies to

Dr Ann Richardson
Independent researcher
e-mail: annrich@resconsult.u-net.com

Methodology Group

Members

Methodology Programme Director
Professor Richard Lilford
Director of Research and Development
NHS Executive – West Midlands, Birmingham

Chair
Professor Martin Buxton
Director, Health Economics Research Group
Brunel University, Uxbridge

Professor Douglas Altman
Professor of Statistics in Medicine
University of Oxford

Dr David Armstrong
Reader in Sociology as Applied to Medicine
King's College, London

Professor Nicholas Black
Professor of Health Services Research
London School of Hygiene & Tropical Medicine

Professor Ann Bowling
Professor of Health Services Research
University College London Medical School

Professor David Chadwick
Professor of Neurology
The Walton Centre for Neurology & Neurosurgery
Liverpool

Dr Mike Clarke
Associate Director (Research)
UK Cochrane Centre, Oxford

Professor Paul Dieppe
Director, MRC Health Services Research Centre
University of Bristol

Professor Michael Drummond
Director, Centre for Health Economics
University of York

Dr Vikki Entwistle
Senior Research Fellow, Health Services Research Unit
University of Aberdeen

Professor Ewan B Ferlie
Professor of Public Services Management
Imperial College, London

Professor Ray Fitzpatrick
Professor of Public Health & Primary Care
University of Oxford

Dr Naomi Fulop
Deputy Director, Service Delivery & Organisation Programme
London School of Hygiene & Tropical Medicine

Mrs Jenny Griffin
Head, Policy Research Programme
Department of Health
London

Professor Jeremy Grimshaw
Programme Director
Health Services Research Unit
University of Aberdeen

Professor Stephen Harrison
Professor of Social Policy
University of Manchester

Mr John Henderson
Economic Advisor
Department of Health, London

Professor Theresa Marteau
Director, Psychology & Genetics Research Group
Guy's, King's & St Thomas's School of Medicine, London

Dr Henry McQuay
Clinical Reader in Pain Relief
University of Oxford

Dr Nick Payne
Consultant Senior Lecturer in Public Health Medicine
ScHARR
University of Sheffield

Professor Joy Townsend
Director, Centre for Research in Primary & Community Care
University of Hertfordshire

Professor Kent Woods
Director, NHS HTA Programme, & Professor of Therapeutics
University of Leicester

continued

HTA Commissioning Board

Members

Programme Director
Professor Kent Woods
Director, NHS HTA Programme, &
Professor of Therapeutics
University of Leicester

Chair
Professor Shah Ebrahim
Professor of Epidemiology of Ageing
University of Bristol

Deputy Chair
Professor Jon Nicholl
Director, Medical Care Research Unit
University of Sheffield

Professor Douglas Altman
Director, ICRF Medical Statistics Group
University of Oxford

Professor John Bond
Director, Centre for Health Services Research
University of Newcastle-upon-Tyne

Ms Christine Clark
Freelance Medical Writer
Bury, Lancs

Professor Martin Eccles
Professor of Clinical Effectiveness
University of Newcastle-upon-Tyne

Dr Andrew Farmer
General Practitioner &
NHS R&D Clinical Scientist
Institute of Health Sciences
University of Oxford

Professor Adrian Grant
Director, Health Services Research Unit
University of Aberdeen

Dr Alastair Gray
Director, Health Economics Research Centre
Institute of Health Sciences
University of Oxford

Professor Mark Haggard
Director, MRC Institute of Hearing Research
University of Nottingham

Professor Jenny Hewison
Senior Lecturer
School of Psychology
University of Leeds

Professor Alison Kitson
Director, Royal College of Nursing Institute, London

Dr Donna Lamping
Head, Health Services Research Unit
London School of Hygiene & Tropical Medicine

Professor David Neal
Professor of Surgery
University of Newcastle-upon-Tyne

Professor Gillian Parker
Nuffield Professor of Community Care
University of Leicester

Dr Tim Peters
Reader in Medical Statistics
University of Bristol

Professor Martin Severs
Professor in Elderly Health Care
University of Portsmouth

Dr Sarah Stewart-Brown
Director, Health Services Research Unit
University of Oxford

Professor Ala Szczepura
Director, Centre for Health Services Studies
University of Warwick

Dr Gillian Vivian
Consultant in Nuclear Medicine & Radiology
Royal Cornwall Hospitals Trust
Truro

Professor Graham Watt
Department of General Practice
University of Glasgow

Dr Jeremy Wyatt
Senior Fellow
Health Knowledge Management Centre
University College London

Current and past membership details of all HTA 'committees' are available from the HTA website (see inside front cover for details)

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.ncchta.org) 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.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.ncchta.org

ISSN 1366-5278