

Managing uncertainty in healthcare

Report of a meeting organised by NICE and AHRQ

Introduction

There are certain challenges that confront virtually all health systems, irrespective of the means by which they are funded and administered. One such is the management of uncertainty: specifically, knowing what to do when data on the effectiveness or the cost-effectiveness of new medicines or procedures is incomplete or inadequate, but decisions have nonetheless to be taken on whether to purchase and supply them to patients. A similar issue arises when there is suspicion that a procedure or medicine already in use may be relatively ineffective or represent poor value for money.

Given the universality of these problems, NICE and AHRQ judged that the topic merited an international workshop. The meeting - held in London on May 22, 2008 - was principally an Anglo-American event, but included individual participants from other countries in Europe and Asia.

The more expensive that medical procedures become, the greater the importance of managing uncertainty of this type. That said, the issue is not a new one, and some medical systems have already responded. “Coverage with evidence development” (CED) is a concept already in use in publicly funded health care schemes in the USA; “Only in research” (OIR) is an option available to NICE appraisal committees when deciding whether something should be made available through the NHS. But experience of their use is still very limited, and there is general agreement that the concepts need to be further refined and developed to be sure they properly serve the purposes for which they are intended.

Aims and objectives

In his role as facilitator, Professor Trevor Sheldon of the University of York outlined the aims and objectives of the day. He began by listing what he saw as the starting assumptions of all participants gathered in the meeting room. These included:

- A desire to use available health care resources in a way that improves the health of the population.
- A wish to see the provision of incentives designed to encourage innovations by people and companies for improving health care.
- A belief that there will usually be uncertainties about cost and effectiveness when making coverage decisions that will determine how public money is spent and on what. Uncertainty means that some decisions may prove not to have been in the public interest.

- A conviction that research is a way of acquiring evidence to reduce uncertainty, and so increasing the probability of making good coverage decisions.
- A belief that because of the close links between research findings and coverage decisions - ie the need to collect the former in order to optimise the latter - the two may need to be considered alongside one another rather than separately.

These assumptions, he said, raise a series of questions that need to be tackled:

- How do we develop a decision or policy framework that encompasses both coverage and research decisions?
- How do we decide if we have enough evidence to make a decision?
- What mechanisms do we need for implementing research decisions, and ensuring that the work is actually done? Could it be through pricing, or will it require some system of rules?
- How will these mechanisms be received by policy makers, industry and the public?
- What are the best methods of generating evidence within this mechanism? How do we make sure we get the right sort of evidence to reduce uncertainty in the right way?

Although these two lists provoked no significant dissent, several participants wished to emphasise particular aspects of them including possible priority conflicts between academics and commissioners, the need to consider existing activities as well as new ones, the role of regulatory authorities in the process, and the importance of reducing inequalities in health – the last of these being, of course, an intention that will vary from country to country.

Why uncertainty matters: economics and citizens

The first session comprised four perspectives on the issue of uncertainty, beginning with that of the University of York economist Professor Karl Claxton. He emphasised that he would be talking about principles not methods – and in particular the principles that underpin the economic assessments that need to be made of a technology if good decisions are to be taken on whether to adopt it. He illustrated his thinking by reference to hypothetical technologies offering various health outcomes at various costs. He reminded his listeners that, within a fixed budget, you must be sure that the health gains you make by spending your money in some new way are not outweighed by the health gains you forgo through not

spending it in the old way. Costs matter because costs are other people's health somewhere else.

Investing in research is important not because it satisfies curiosity, but because it creates the possibility of making decisions that maximise the cost effectiveness of your spending. Evidence, in this sense, can be said to have a health value of its own. Thinking in these terms makes it easier to decide when it is worth spending money to gain extra evidence, and when it is not. The grey area – the area that constitutes a justification for OIR/CED type decisions – is characterised by evidence for or against effectiveness (or cost-effectiveness) that is weak or inconclusive. But this still leaves a raft of questions still to be answered. What kind of evidence? Who should acquire it and how? What incentives should there be for its collection? Who should pay? Could the manufacturer be persuaded to set a lower price that makes the technology cost-effective anyway? One of the key conclusions was that early evaluative research is valuable, and delaying research imposes its own costs.

The next presentation - a lay perspective - was from Mr Paddy Storrie, until recently a member of the NICE Citizens' Council which, as he explained, comprises some 30 lay people chosen to represent a rough cross-section of the population. Set up to advise and comment on various technical and ethical issues facing NICE, its January 2007 meeting was devoted to considering the "Only in research" option available to NICE appraisal committees. This was the most technical issue the Council had confronted and, like all its reports, represented an informed public view, but not necessarily the public view.

The Council's discussion showed that most people had a mature acceptance of uncertainty in medical treatment and of particular events such as side effects which were, of their nature, unpredictable. They were also tolerant of acting as "guinea pigs" where there was real uncertainty, and were prepared to contribute to knowledge by participation in clinical trials.

They were less forgiving of flawed treatments - such as the prescription of steroids for head injured patients - that had survived in routine practice because they had never been properly evaluated. There had been unanimous approval for the OIR principle. But this principle had to be strongly enforced. If a treatment were to be made available with only vague or unenforceable suggestions that further evidence should be collected, the Council feared that it might be disregarded.

Among the Council's key provisos was one concerning treatments for a terminal condition where the treatment in question was the only one available. The Council felt that under these circumstances NICE should take a more lenient view – which might include opting for an OIR decision when it might otherwise have chosen simply to say no. Another proviso was the need for realism when insisting that patients be entered into clinical trials; the facts of geography mean that this will often be easier for those living in cities rather than in rural areas.

Among the Council's key concerns were that an OIR recommendation should not be a refusal in disguise; and that NICE must resist media, patient, industry or other outside pressures and sustain its robust judgements. Facing up to the media in particular could be a problem. Bayesian statistics cannot compete with a photo of a sick child.

Why uncertainty matters: American perspectives

Alan Garber of Stanford's Centre for Primary Care and Outcomes Research began his presentation on some of the practicalities of decision-making by pointing out that no US payer employs a process in which cost-effectiveness is used to determine coverage. And even national programmes such as Medicare are contracted out to local suppliers who make their own decisions. His own experience comes from years of the medical advisory panel of Blue Cross/Blue Shield. The panel's criteria for using a technology are that it has government regulatory approval and adequate evidence of effectiveness, that it must improve health outcome, be as beneficial as any established alternatives, and yield improvements outside purely investigational settings.

The middle three of these five criteria are the ones that matter. And they can be problematic. The Medicare programme, for example, is governed by rules that say "Medicare shall not pay for anything that is not reasonable and necessary for the treatment of a condition". But what does "reasonable and necessary" mean? Moreover in practice the negative "it shall not pay if it is not reasonable and necessary" has flipped into "must pay if it is reasonable and necessary".

Alan Garber also talked of the difficulty of deciding what constitutes sufficient evidence of benefit. Does it have to be an RCT, or would observational data or a registry of some kind be sufficient? He also spoke of the extent to which backtracking from a previous decision can undermine an organisation's credibility. Using the examples of high dose chemotherapy for breast cancer and lung volume reduction surgery for emphysema - therapies which have not lived up to the expectations

derived from earlier observational studies - he illustrated some of the political and other pressures brought to bear on those who have to make coverage decisions. He concluded that some of the biggest mistakes have been a result of relying on observational data alone, and commented that a key question is knowing when it is necessary to insist on RCT data. He also wondered if, in the future, it might be possible to get “something close to randomisation without actually doing it”.

The final presentation of the session - on dealing with uncertainty in Medicare coverage decisions - was given by Sean Tunis of the US Center for Medical Technology Policy. He too referred to the Medicare definition of what the programme will pay for: "adequate evidence to conclude that the item or service improves the health outcome". This may sound cut and dried and scientific, but the concept of adequacy is a judgment, not a scientific notion. Talking from personal experience he said that evidence is almost never “adequate”.

He talked of his experience in trying a variety of measures to reduce uncertainty. These ranged from providing cover but urging better studies through to restricting coverage to centres that collected data (poor man’s CED, as he dubbed it). In the end he and colleagues favoured full blown CED: payment only if all patients are enrolled in an appropriately designed study. This has been applied in various cases including lung volume reduction surgery, FDG-PET in oncology and for suspected dementia, and the off-label use of drugs for colorectal cancer. In most cases, according to Sean Tunis, these studies have proved to be “informative failures”. None has produced the quality of data really needed to make future coverage decisions, although NETT (the National Emphysema Treatment Trial of lung volume reduction surgery) did provide findings that essentially eliminated the use of the procedure in the US.

In spite of this rather dismal account, Tunis remains optimistic that CED can be made to work, provided a series of challenges can be met. These include: the selection of the right topics, and at an early stage in their development before it is too late to start contemplating CED; a reliable source of funds; and a strategy for carrying out fast, simple and cheap studies.

Questions from the audience included the possibility of international co-operation and the issue of what guidance should be given about the kind of evidence that decision making bodies might require. Tunis talked of developing a library of documents spelling out matters such as types of

patient and required outcomes. Clearer guidance from payers and the medical professional societies and others on key design elements of proposed trials might help to generate the right kind of evidence.

CED in the US

Following the mid-morning break Jean Slutsky, director of the Center for Outcomes and Evidence, presented her thoughts on public and private payer experience with conditional reimbursement in the USA. She believes there is a widespread sympathy for the idea. She described the US “landscape” as well intentioned, ad hoc, and relatively (with some exceptions) uncontroversial. But conditional reimbursement has proved easier to accept as a concept than to apply in reality, and has so far been characterised by some disappointments as well as some stunning successes.

Conditional reimbursement, she said, can be a good thing if fair (to everyone, including patients) and well-designed. But there is no systematic process for implementing it. By and large, opportunities are simply seized as they arise. Once a conditional coverage decision has been made, finding a payer to cover the cost of the research component is hard. This can favour less rigorous and less innovative study designs.

As far as the future is concerned, Slutsky sees private-public funding and participation as likely to be essential. More effort will be required if conditional reimbursement study designs and protocols are to improve, and there is a continuing need to tackle issues such as ethics, priority setting, and knowing when the accumulated evidence has reached the point where CED can justifiably give way to unconditional coverage.

None of the five speakers during this first batch of the morning’s presentations suggested that variants of the OIR/CED principle were as yet satisfactory, and all in various ways pointed up the difficulties of devising effective and workable improvements. But all endorsed the need for such schemes, and none suggested that the goal of reducing uncertainty was unattainable.

Frontline experience, American and British

The next five presentations offered practical illustrations of various research projects that have been undertaken to inform healthcare decision making. First to speak was Tamara Syrek Jensen of the US Centers for Medicare and Medicaid Services. Following initial difficulty in selling the idea of CED internally and to the public, her perception is that it’s now quite popular and accepted in the US. To serve the needs of

transparency, the Centers have compiled a guide to CED which explains that the scheme comprises a registry component and a research study component, and describes the basis on which each operates.

Unfortunately, CED has now outgrown some of the criteria and the rules laid down in the document – which, as a consequence, is actually holding it back. Taking a decision to rewrite the document is not, apparently, as straightforward as it might seem. A further complication is that although the guidance was devised nationally, most actions and decisions are taken at local level. There are also various statutory and regulatory provisions that hamper change.

She referred specifically to one example which she thought - though she added a question mark - could be defined as a success. This was the National Oncologic PET Registry (NOPR), a study of 23,000 patients who'd been treated in 1200 facilities nationwide. The data are now being scrutinised, and in due course Medicare will decide whether or not to give this technology coverage without limitation. The challenge created by this study is that no threshold has yet been defined for taking CED out of the equation, and going to outright coverage.

Issues raised by the audience following the presentation included: the extent to which certain issues might be more swiftly resolved through international studies; whether CED really does have the widespread support claimed for it and how far this support would survive significant constraints on demand; whether patients ever refuse treatment if they are required to take part in a trial; and whether many medical professionals were that much better informed about trial methodology than lay people.

Next to speak was Barney Reeves of Bristol Royal Infirmary who described the UK's verteporfin photodynamic therapy (PDT) cohort study. For certain forms of macular degeneration this treatment has received an OIR recommendation. NICE had felt, on the basis of an existing trial, that there were uncertainties about cost-effectiveness in some sub-groups of patients – though cynics have suggested that it was really to allow more patients to receive the treatment while preventing a free-for-all. Reeves described how he and his colleagues had set up an OIR treatment register intended to answer a series of questions including whether PDT is provided within the NHS in the same manner that it was within trials, and how effective and cost-effective it is.

Summarising his experience he said that the study was extremely difficult to do, and he wasn't entirely happy with the data finally generated, partly because the scheme's organisers had been too ambitious. Personally he

believes that treatment registers need clear and very limited objectives if they're to succeed, and compliance with them has to be mandatory: no data, no payment should be the rule. He now feels that the money spent would have been better used running a pragmatic trial. But the PDT cohort study has nonetheless provided some valuable data, including the relationship between visual acuity and quality of life.

The third presentation of this session was an overview of the biologics register set up eight years ago by the British Society for Rheumatology and intended to monitor the long term safety of several anti-TNF agents in the treatment of severe rheumatoid arthritis. The manager of the register, Mervyn Hogg, described how it aims to compare cohorts of 4,000 patients treated using biologics with another 4,000 being treated by more conventional means. Funded by the drug industry, the register was set up in response to a NICE recommendation. Having described how the system operates, and outlined the safeguards in place to ensure that the industry does not exert inappropriate influence, Mervyn Hogg reported that most patient cohorts had now been filled, and were into the five year follow up phase.

The apparent success of the register has prompted interest in setting up other such studies in psoriasis and ankylosing spondylitis.

Following the presentation, a member of the audience commented on the extent to which the direction taken by a registry is determined by whoever is organising and funding it, and suggested that those set up by professional societies would not necessarily want to collect the kind of data best suited to serving the concerns of the people gathered in the meeting room. Mervyn Hogg agreed. Discussion at the planning stage was the way to deal with this, he thought. Another questioner wondered about possible conflicts if the results, when available, were not to the liking of all concerned with the project, especially the drug companies. And another expressed scepticism about the likelihood that all data on adverse events would actually find their way into the register. His own experience suggested to him that they didn't.

A presentation from surgeon Tom Treasure of University College, London provided a further example of British experience, this time in the assessment of an existing practice. While surgery for lung and breast tumours has become less frequent and/or less radical over the years, this has not been so much the case with colorectal cancer. Radical surgery is clearly indicated for the primary tumour in the colon, because the tumour might cause an obstruction, but not necessarily for secondary metastases.

Little attempt has been made to compare the benefits of operating on metastases versus not doing so - this in spite of a series of papers that have shown how measurement of a circulating marker, CEA, is a predictor of survival. There is also a lack of data on patients' quality of life and state of mind following surgery to remove metastases, and there is no reference group to suggest what would have happened in the absence of surgery. In short, no satisfactory attempt has been made to decide who should receive surgery, and what benefit it might confer.

The problem, Tom Treasure suggested, is deep rooted, and starts with the false assumption that all colo-rectal cancers are much the same. They are not. Longer survival following surgery is, he suspects, as likely to be associated with selection for surgery as with the benefit of the operation itself. For all these reasons a trial is needed, and there is now backing for one. Called PulMICC it has already registered nearly 100 patients and shown - in the face of scepticism - that randomisation is possible.

The presentation was followed by questions and discussion on the impetus for organising work of this kind, and the extent to which clinicians themselves can be relied upon to drive it, and what role their professional associations might play in encouraging it. It was noted that neonatologists in the UK had voluntarily decided collectively to use some new treatments (extra-corporeal membrane oxygenation, and cooling, for example) only in the context of randomised trials until more was known about their effects. Examples of work of this kind - and others intended to find out if there is a case for disinvestment in a treatment - are few in number; but one participant did refer to an MRC funded series of trials of palliative radiotherapy and lung cancer. This demonstrated shorter courses of radiotherapy to be just as effective as longer courses. Practice in Britain has now changed - though not, it seems, elsewhere.

Frontline experience, Canadian and European

The penultimate presentation switched the geographical focus back to North America - this time to Canada, and to something of a success story. Leslie Levin is professor of medicine at the University of Toronto and head of Ontario's Medical Advisory Secretariat. In this latter capacity he spoke about the province's Conditionally Funded Field Evaluation (CFFE) scheme, a form of CED. Having described how CFFE is managed and funded, he listed the ten studies completed. These included a safety evaluation of CT and MRI, and the use of endovascular repair for abdominal aortic aneurysms. The results of each study, he added, have either impacted on policy and decision making, or are in process of so

doing. One study of drug-eluting stents that covered 21,000 patients over 18 months produced findings which contradicted those of the published RCTs.

He also listed 24 more studies that are ongoing, including some on PET scanning – a procedure that Ontario will not make an insured service until its clinical utility is better understood. He admitted that this particular work had generated considerable flak from the public and from nuclear physicians. Other studies included the safety of smart infusion pumps and the use of neuromodulation for depression. The CFFE list includes some technologies that are already in routine use. The methodology for all these projects is determined by the types of question being asked, and ranges from RCTs and registries to observational studies.

In response to an audience question inquiring if a system of this kind might be unique to the Canadian arrangements for financing healthcare, Professor Levin pointed out that the system is not hugely expensive. The 34 studies had together cost around \$8.5 million for the data collection and analysis. Another questioner commented on the fact that most of the topics chosen for analysis were individual therapies. Professor Levin conceded that they had indeed focussed until recently on single technologies, but more recently they were beginning to look at the integrated technology surrounding one particular health condition, diabetes.

The final speaker, Sun-Hae Lee-Robin, talked about the work of the European Network for Health Technology Assessment, Eunetha. More specifically she focussed on its work package 7 which is devoted to monitoring the development of new and emerging technologies, and facilitating the generation of new evidence.

Many countries in Europe undertake such monitoring, but at many different levels of detail, and with varying degrees of intervention. Some would like to share information and co-ordinate their efforts, particularly in the case of rare conditions. Eunetha is aiming to accommodate the desire by developing tools that will allow this to be done. Technologies identified by partners as of particular interest include percutaneous aortic valve replacement, coronary CT scanning, HPV vaccination and telemedicine.

Following this presentation, a member of the audience pointed out that, when assessing the virtues of a particular procedure, registers can be attractive because those who set them up usually think the register will

support the original decision to make at least limited use of the technology. What happens if the register data go against the original intuitive judgment? Sun-Hae Lee-Robin admitted - to laughter - that she had no idea.

Another participant pointed out that assessments of an existing technology is usually seen in terms of making a case for disinvestment. By contrast, he recounted the tale of an unlicensed drug of no commercial interest that some clinicians thought might be useful for reducing episodes of apnoea in premature babies. A neonatologist managed to organise a trial and found that the agent concerned could reduce cerebral palsy and developmental delay. If it had been properly studied 30 years previously, its true value would have been appreciated and exploited much earlier. One of the earlier speakers, Alan Garber, said that this anecdote prompted him to comment on something he feared might be a more common occurrence: not people discovering something new and unexpected from register data, but instead confirming their prejudices by managing to find precisely what they were looking for. Many participants, however, remained committed to the value of register data.

Another comment concerned the exploitation of the new NHS information infrastructure Connecting for Health. This system, properly used, might offer an alternative to creating a plethora of registries. Not all the UK participants seemed convinced.

Strategic priorities

The first session of the afternoon was to have been organised around small groups in which participants would compile lists of what they felt, in their experience, were the barriers, challenges and obstacles to reducing uncertainty. However, Trevor Sheldon decided that the questioning and discussion of the morning had thrown up most of what concerned participants and proposed that the rest of the time be spent on discussing ways forward. So instead he presented, for their approval, a list of challenges that he himself had drawn up based on what had been said.

1) The development of an overall conceptual framework for making decisions on coverage. This must encompass four sorts of decision. There is technology that comes with a lot of data generated by industry or academia and to which it's possible to say, yes, the evidence justifies using the procedure; or yes (make it available to all), but only on condition that still more research is done. The latter he called "yes +". Or you can say no based on the current research, or no but generate more

research because it's worth doing (a "no +") (and so only make it available to some people participating in the research).

2) The need to prioritise. If resources are constrained, there has to be a method of generating priorities (for both the technology and the further research to be carried out), an agreed list of who should be consulted (public, professionals, policy makers etc), and some thought about what happens to procedures not chosen.

3) Ensuring that the prioritised research is actually conducted. This means that sufficient resources must be available, and that international collaboration - where this might prove helpful - has been arranged.

4) Ensuring that the evidence generated is fit for purpose. That is, it will sufficiently reduce uncertainty and be suitable for influencing policy and practice. This means it must be relevant, timely and cost-effective.

5) The process must be legally, ethically and otherwise acceptable to the public, government, industry, academics and the media. This will make it more likely to be happen, and more likely that coverage decisions will be accepted.

6) There must be appropriate incentives to ensure that industry and other parties involved make the process happen in a timely fashion.

7) It must be clearly decided who is to do what, and how the various elements of the process are to be integrated.

One participant suggested that there should be an eighth challenge: harmonisation.

Following a brief discussion of the implications of some the challenges listed by the Professor Sheldon, the participants broke into four groups to discuss some of the challenges in more detail, and to identify ways of tackling them.

Feedback from groups

The final session of the day began with representatives of the each of the four groups offering feedback on their discussions. It soon became clear that many of these discussions had been free ranging, and bore little (or even no) direct relationship to the issues as identified and categorised by the meeting facilitator.

Group 1 claimed to have reached no consensus on anything, including the four yes/no category distinctions outlined in Trevor Sheldon's first challenge. Some members felt that the plus in "yes+" decisions might in practice tend to be ignored. Its utilisation, they argued, would certainly depend on research being embedded in clinical practice. The group also talked about prioritisation, but then focussed on something not in the question as posed: how interventions get into the frame in the first place. Some participants thought that because the number of interventions dubbed 'only in research' was never likely to be great, there should be no difficulty in funding the required work.

On the matter of ensuring that research is actually done - challenge three - flagging up research recommendations early on would surely help, as would spelling out more concisely the research questions to be answered. More liaison between NICE-like bodies, researchers, funders and service commissioners - all with different agendas - would also be valuable.

The group discussed the obstacles created by bureaucracy, and also intervention-based registers which, they felt, might have a place in reducing uncertainty about matters such as complication rates and 'indication creep'. But such uses would have to be clearly defined and might even then prove quite limited.

Group 2 reported that their discussion - "productive" they said - had begun by considering the firm criteria required for making decisions such as cost effectiveness in relation to a fixed threshold, and then moved to pondering less defined features of the process such as judgements of social value. A discussion of the Trevor Sheldon's four category distinctions brought them to the conclusion that this arrangement only makes sense if the process of making decisions is not fractured; in other words, that an "only in research" decision will probably become simply a soft "no" unless it is directly linked to the prospects of actually getting that research done.

They also offered an example of the practical difficulties that can be thrown up in decision making. If you think something new is probably worth having, but the definitive trial is going to report in a week's time and you know you can't change your decision, you're clearly going to delay making it. But what if it's in a month's time, or six months, or five years?

On the matter of ensuring that research is actually conducted, the group felt one could be reasonably confident that it would go ahead following a

“no +”, not least because of the powerful financial incentive: if you want access to the market, you do research. There was less confidence that it would follow a “yes+”, at least without a serious penalty or reward or a strict contract of some kind. Where there is good routine collection of registry type data, “yes+” can be done cheaply.

They also discussed acceptability to the public and the media – concluding that the latter, in spite of the periodic slings and arrows they fire off, are an essential element of accountability.

Group 3 discussed timeliness: important because undue delays may render CED or OIR research findings irrelevant. One way of achieving this is through incentives; another is by allying the payer more closely with whoever is generating the data. They also discussed the kinds of incentives - including patent length - that might be appropriate.

A more radical suggestion was to change the law on licensing so that payers have access to information besides efficacy and safety. More comparative data would be valuable. Legislation could include the concept of added value: a new drug should have something that an existing one doesn't. A right to have access to all the clinical trials that have been done would also be useful. And if a drug does not work, maybe the industry should repay the money.

Group 4 discussed the possibility of harmonisation across jurisdictions, and concluded that it would be difficult if not impossible to reach agreement. They also discussed the importance of communication between companies, government agencies and researchers. Having the right kind of communication and having it early were both seen as important. But there was concern that government officials might feel bound by discussions, and treat them as contractual obligations.

There is also the issue of whether discussions are held before or after any regulatory approval that may be required. In this context the group also discussed reference pricing in which companies are encouraged to provide comparative information in the knowledge that evidence of efficacy greater than that of a competing product will earn them a return higher than is being paid for this product.

On the question of paying for research the group considered the option of imposing a tax of some kind on drug companies or device manufacturers – but reached no definite conclusion, except to emphasise that a dedicated budget from some source is essential.

Educating the media will be an important feature of any successful scheme, and the group was torn between cynicism and idealism. Either way, if the media are to be educated, it will best be done at the beginning of the process – the aim being to ensure they understand why the process is being undertaken. But the task is a hard one.

More generally an agreed process has to be in place from the outset, defined perhaps by an expert committee convened to spell out the kind of evidence that would be required, and decide if the work is likely to be expensive.

Final thoughts

With the completion of the feedback session, facilitator Trevor Sheldon called for any final comments or clarifications. One participant wondered if there shouldn't have been more discussion of the extent to which anyone could be satisfied with anything less than RCT evidence. The problem, of course, is that with so much needing to be learned it is not possible to hope for RCT evidence on every issue. Another participant suggested that an insistence on RCTs would spell death for post-marketing surveillance. There was also discussion of the role of pragmatic studies and of registers – the issue in the second case being whether ambitious disease-based registers covering every presenting patient are essential, or whether more modest intervention registers are acceptable.

The final task of the meeting was to decide how best to take the discussion forward in the future. There was a general feeling that the meeting had been useful – so Professor Sheldon sought suggestions from participants on which topics merited more sustained thought and discussion by an international group of this kind. The list included:

- The value of international collaboration, and an attempt to clarify nomenclature.
- The role of observational and other study techniques.
- Detailed critiques of specific examples of current practice to illustrate possible methodologies.
- More involvement of service commissioners.
- Value of information analysis and its role in priority setting.
- More consensus about the outcomes that are important.
- Financial incentives for research, and who will fund it

A show of hands revealed an appetite for future meetings - perhaps preceded by the production of reports or papers on particular issues - at which progress in managing uncertainty could be discussed and compared. The presence of service commissioners at such a meeting would greatly enhance its value.