ELSEVIER

Social Science & Medicine 60 (2005) 937–947

SOCIAL SCIENCE & MEDICINE

www.elsevier.com/locate/socscimed

Assessing the benefits of health research: lessons from research into the use of antenatal corticosteroids for the prevention of neonatal respiratory distress syndrome

Steve Hanney[a,*], Miranda Mugford[b], Jonathan Grant[c], Martin Buxton[a]

[a]*Health Economics Research Group, Brunel University, Uxbridge, Middlesex UB8 3PH, UK*
[b]*School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, UK*
[c]*RAND Europe, Cambridge, UK*

Available online 18 August 2004

Abstract

Do the benefits from health research justify the resources devoted to it? Addressing this should not only meet increasing accountability demands, but could also enhance understanding of research utilisation and how best to organise health research systems to increase the benefits. The process from basic research to eventual application and patient benefit is usually complex. The use of antenatal corticosteroids when preterm delivery is expected has featured large in the debates about research utilisation and provides an insight into these complexities. Based on an analysis of previous modelling of research utilisation and payback assessment, a framework is developed in which the existing literature on the use of corticosteroids, combined with new material developed by the authors, can be reviewed and synthesised. The move from animal studies to human trials was undertaken by the same individual. Some early clinical application of the findings occurred concurrently with a series of further trials. Nevertheless, the implementation of these findings stalled rather than accelerated as is predicted by some models. The eventual systematic review of the trials played a part in the development of the Cochrane Collaboration and increased the impact on practice. Further implementation approaches were used in various countries, including clinical guidelines, a National Institutes of Health Consensus Conference, and various implementation projects within the UK. This paper shows how an assessment of the benefits from this stream of research and utilisation projects can be constructed. It concludes that the application of a model for assessing payback can help to demonstrate the benefits from the research in this field and enhance our understanding of research utilisation.

Keywords: Assessing research benefits; Research payback; Research utilisation; Respiratory distress syndrome; Corticosteroids

Introduction

Do the benefits from health research justify the resources devoted to it? This question is increasingly asked, especially when the funding could otherwise be spent directly on providing health care. Various approaches to, and reasons for, assessing the benefits from health research have been advanced (Drummond, Davies, & Ferris, 1992; National Institutes of Health ((NIH), 1993; Buxton & Hanney, 1996; Grant, Cottrell, Cluzeau, & Fawcett, 2000; Smith, 2001; Croxson, Hanney, & Buxton, 2001). Some reasons revolve around

*Corresponding author. Tel.: +44-1895-265444; fax: +44-1895-269708.
E-mail address: stephen.hanney@brunel.ac.uk (S. Hanney).

0277-9536/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.socscimed.2004.06.038

the increasing demands for accountability for existing research funds, the desire to provide justification for current levels of expenditure and advocacy for extended funding. Others relate to the potential use of assessments in helping both to increase understanding of the processes involved and assist identification of how research systems can best be organised to enhance utilisation and benefits, especially for patients. Finally, it is claimed that assessment of utilisation could provide incentives for greater attention to be given to activities aimed at enhancing utilisation.

Previous modelling of research utilisation and payback assessment recognises that only sometimes do flows of health research knowledge make a direct impact on policy and practice: more often they simply add to the pool of knowledge (Kogan & Henkel, 1983; Hanney, Packwood, & Buxton, 2000). This increases the difficulties of assessing the benefits from research. Not only do the types of benefits that can flow from research have to be clarified, and methods selected for assessing whether these paybacks are accruing, but it is also important to consider how to identify which research is responsible for any payback achieved.

Few topics have figured so prominently in the research utilisation debate as the use of antenatal corticosteroids to prevent neonatal respiratory distress syndrome (RDS) when preterm delivery is expected. Indeed, this issue became a *cause celebre* among those concerned with encouraging greater research utilisation (Department of Health, 1993; Haines & Jones, 1994).

Models of research utilisation and payback assessment

We draw on models of research utilisation and frameworks for payback analysis in order to present and organise a wide range of available evidence. Theories of diffusion of innovations generally examine patterns of adoption of new findings. For example, Rogers (1995) developed the concept of the S-shaped adoption curve which helps inform analysis of how far uptake of research findings will occur 'spontaneously' and how far specific steps are necessary to encourage implementation. Rogers' S-shaped curve shows the number of adopters rising slowly at first, then accelerating and finally increasing at a gradually slower rate as fewer and fewer remaining individuals adopt the innovation. The part of the diffusion curve from about 10% adoption to 20% adoption, he states: 'is the heart of the diffusion process. After that point, it is often impossible to stop the diffusion of a new idea, even if one wished to do so' (Rogers, 1995, p. 259). Haynes and Haines (1998) describe a path from evidence generation to clinical application that can involve a series of stages, including synthesising the evidence through systematic reviews and formulating clinical policies. Such models of research utilisation adopt the approach of working forwards from the production of new evidence and examine its implementation. Not all models of payback assessment work in this direction. In an attempt to develop a more systematic approach than evidenced by previous anecdotes, Comroe and Dripps (1976) identified key aspects of then current clinical practice in the cardiovascular field and attempted to work backwards to locate the crucial bodies of knowledge behind them. They showed that much of the key research was not clinically oriented when it was undertaken, although the replicability of Comroe and Dripps' work has been challenged (Smith, 1987).

Most models of payback assessment from the health economics literature focus primarily on providing a quantitative ex-ante assessment of the likely magnitude or value of the payback from research that is being considered for funding. Townsend, Buxton, and Harper (2003) reviewed the various models designed to be used in this way. Economic evaluations can be undertaken at various times in the development of a stream of research and play various roles (Sculpher, Drummond, & Buxton, 1997), sometimes forming a key part of payback assessments. In an ex-post assessment of payback linked to the NIH (1993), Drummond et al. (1992) assessed the payback once a body of research on diabetic retinopathy had been completed but on the basis of expert opinion about likely levels of utilisation rather than data on actual uptake.

Our approach, though informed by various quantitative models, was developed to focus on actual take-up levels and aims to enhance understanding of the processes and linkages that connect payback to the original research. Our framework for assessing payback consists of two elements: a multidimensional categorisation of benefits (ranging from knowledge production, through an improved information base for policymaking, to the final outcomes of health gain and broader economic benefits) and a model of how to apply this categorisation (Buxton & Hanney, 1996; Hanney et al., 2000).

The model contains a series of stages, but key features are the interfaces, and associated levels of permeability, between research and the wider professional and political environments that constitute the context. The initial interface involves the specification of research to meet identified needs. The inputs into research projects, and the subsequent processes, lead to the primary outputs: the production of knowledge in the form of publications and, quite often, capacity building for future research. Then there is the dissemination interface, at which point the research findings usually enter the pool of knowledge—from where they often feedback into further research. The findings might also be disseminated to the wider society of industry, policymakers, practitioners, and members of the

public—especially relevant patient groups (Hanney, Gonzalez–Block, Buxton, & Kogan 2003a).

The next stage for the assessment of benefits focuses on the generation of 'secondary outputs' in the form of research informed policies and products. Such policies can range from national public policies to local administrative decisions and clinical guidelines developed by professional groups (Hanney et al., 2003a). Policies are based on many factors, sometimes including systematic reviews of all relevant, rigorous research. There are various ways in which practitioners can be encouraged to adopt or apply research findings and/or research informed policies (Grimshaw et al., 2001). Application of the research findings by practitioners should lead to the final outcomes, in the form of the benefits to the health and economic sectors, including health gains, cost savings and a healthy workforce.

This model was originally developed for health services research (Buxton & Hanney, 1996). When the focus is on more basic research it is likely that greater emphasis will be given to the various phases of knowledge production. With some adjustments, however, the framework set out in the preceding paragraphs seemed appropriate for the current study.

Methods for studying payback

For the analysis described here we were able to draw on the considerable existing literature about the use of antenatal corticosteroids. Some of this had been undertaken, in previous studies, by the authors of this article: detailed economic and payback analysis (Mugford, Piercy, & Chalmers, 1991; Mugford, 1993); key informant interviews (Hanney, 1994); surveys of potential users (Hanney, Soper, & Buxton, 2003b); and bibliometric analysis (Grant, Green, & Mason, 2003). The latter analysis involved an attempted replication of Comroe and Dripps' study in which the Wellcome Trust's Policy Unit worked backwards from major clinical advances in neonatology, including the use of corticosteroids when preterm delivery is expected (Grant et al., 2003). Therefore, even though difficulties with replicating the methods used by Comroe and Dripps mean that the approach of working backwards seems unlikely to be applicable on a regular basis, some key background findings from the Policy Unit's study are relevant here.

For the current exercise, the main methods included: a literature review to identify additional and recent accounts, desk analysis of the collated literature, citation analysis, and, following Antman, Lau, Kupelnick, Mosteller, and Chalmers (1992) and Graham (1997), a review of textbooks—in this case those series of obstetrics textbooks from the 1970s onwards that were available from the London library of the Royal College of Obstetrics and Gynaecology (RCOG). The key feature for our study is the application of the framework described above to the large amount of material gathered and the organisation of that material to produce a comprehensive analysis of payback and the reasons for it, which had not previously been attempted.

This article organises key elements from the existing literature on RDS, plus the new data we gathered, in two ways. First, the material is presented in the chronology of selected principal events displayed in Table 1. Second, the analytical sections below broadly reflect the various steps in the models of research utilisation and payback assessment: primary research; secondary research; recommendations for use including development of policy guidelines and other secondary outputs; attempts to encourage adoption of the research findings and assessments of the degree of research use; and assessment of the benefits from implementation, including the final outcomes.

Overall, while the focus, especially in the early part of the paper, is about the international development of the science, much of our account of the implementation and payback assessment concentrates on developments affecting the use of corticosteroids in two contrasting healthcare systems: UK and USA.

The development of the primary science

The identification of a specific starting point for an exercise such as this is often problematic but here there are strong arguments for starting with the work of Liggins (Liggins, 1969; Liggins & Howie, 1972). At the end of the 1960s he examined how, when glucocorticoids triggered the onset of labour in pregnant sheep, the lambs born prematurely had well-aerated lungs, while many of the control animals died of RDS (Liggins,1969). Further experiments using a range of animals were continued by others in the early 1970s (Avery, 1975).

Human prematurity was a problem gaining increased attention in the 1960s and the sheep model was being researched by various teams. In relation to the use of corticosteroids, however, the contribution of Liggins was particularly important because he also conducted the first trials in humans. Just 3 years after the influential animal study he published a randomised controlled trial (RCT) of betamethasome therapy involving 282 mothers in whom premature delivery was expected. The aim was to reduce 'the incidence of neonatal respiratory distress syndrome by accelerating functional maturation of the fetal lung' (Liggins & Howie, 1972, p. 515) and the results provided 'sufficient evidence of beneficial effects on lung function and of absence of adverse effects to justify further trials' (p. 524). Their studies continued and expanded; the eventual numbers reported from their

Table 1
A chronological table of selected events

Year (s)	Event
1969	Publication of Liggins' animal research (Liggins, 1969)
1970–1973	Continuing research on acceleration of animal lung maturation by glucocorticoids—reviewed by Avery (1975)
1972	Publication of first trial on humans (Liggins & Howie, 1972)
1974–1984	NIH conference leads to Collaborative Group research (Collaborative Group, 1981, 1984)
1977–1979	Publication of the initial subsequent five clinical trials on humans included in Crowley (1989) systematic review
1979	Strong recommendation for use in a review in *The Lancet* (Ritchie & McClure, 1979)
1980	UK survey of self-reported behaviour suggests 82% of Member or Fellows of the RCOG use it frequently or sometimes (Lewis et al., 1980)
1980–1989	Publication of six further RCTs included in the original systematic review
1984	Editorial in *The Journal of Pediatrics* advocates use (Avery, 1984) following publication of Collaborative Group's findings
1986	Attempt by NIH to show the payback from their trial (NIH, 1993)
1987	Royal Australian College of Obstetricians and Gynaecologists: 76% would use it (Quinlivan et al., 1998)
1989	Systematic review published (Crowley, 1989—also 1990)
1988–1995	Six further RCTs published and included in later systematic reviews
1991	Publication in the UK of economic evaluation of benefits to UK of implementing findings from systematic review (Mugford et al., 1991)
Early 1990s	Level of uptake in England and Wales between15% and 20% (DH, 1993). At the 20% level about 150 neonatal deaths averted annually (Mugford, 1993)
From 1992	Cochrane Collaboration: explanation of its logo that illustrates a review of steroid trials (Cochrane Collaboration, 2003)
1992–1993	Publication of guidelines and other recommendations: BAPM/RCP (1992); RCOG (1992); Executive Letter from the NHS Management Executive (1993)
1994	Low uptake in USA results in: NIH Consensus Conference (1995); a further systematic review (Crowley, 1995); and a project implementing conference recommendations (Leviton et al., 1999)
1996	Last major update of Cochrane systematic review (Crowley, 2002)
1997	Virtually all units in the UK would give corticosteroids to women at risk of preterm delivery (Brocklehurst et al., 1999)
1999	Implementation trial in USA shows increased use from 33% to 58% following traditional dissemination of NIH recommendations and from 33% to 68% following active dissemination (Leviton et al., 1999)

trial made it the largest in the systematic review described below, accounting for about one-third of the more than 3000 women included in the 12 trials (Crowley, Chalmers, & Keirse, 1990).

In the 1970s further RCTs were initiated. Following an NIH workshop in 1974 a large-scale collaborative study, including long-term follow-ups, started in 1976 and reported in the 1980s (Collaborative Group on Antenatal Steroid Therapy, 1981, 1984). The 1981 paper, according to an editorial in the same edition of the *American Journal of Obstetrics and Gynecology*, proved the efficacy 'under certain conditions; however, corticosteroids should be used with caution' (Little, 1981, p. 287). In the 1984 account of the Collaborative Group's follow-up studies the evidence in favour of using corticosteroids was even stronger.

Secondary research

As early as 1981 Crowley undertook a meta-analysis of these trials but this review was first published in a structured form in the *Oxford Database of Perinatal Trials* and *Effective Care in Pregnancy and Childbirth* in 1989 and published in a journal a year later (Crowley et al., 1990). Just 12 trials, out of 23 examined, met the predefined criteria of research quality necessary for inclusion in the systematic review. The results were clear

> Data from 12 controlled trials, involving over 3,000 participants, show that corticosteroids reduce the occurrence of respiratory distress syndrome overall and in all the subgroups of trial participants that we examined.… There is no strong evidence suggesting adverse effects of corticosteroids (Crowley et al., 1990, p. 11).

Overall, the review showed that the reduction in the odds of neonatal respiratory morbidity is about 40–60% and the reduction in early neonatal deaths in babies at risk of RDS is between 25% and 50%. Having reviewed the trials, Crowley et al. went on in the article to report claims that application of the therapy should reduce the costs of hospital neonatal care; they quoted an estimate

suggesting possible savings of $35 million per year in intensive care costs in the USA (Avery, 1984). Following a suggestion from Chalmers, one of the authors of the systematic review, Mugford and colleagues assessed the economic impact that would result from implementation of the findings of the systematic review in England and Wales (Mugford et al., 1991). This indicated a potential reduction in NHS neonatal costs of around £8 million.

Concern about limited implementation of the findings in the USA, despite the strength of the evidence, led the NIH and its affiliate, the National Institute of Child Health and Human Development, to develop a Consensus Conference for which Crowley was invited to update the systematic review. It was subsequently published (Crowley, 1995). Her systematic review on the Cochrane Database was substantially updated in 1996 and amended in 1999 (Crowley, 2002).

Recommendations for use

Recommendations for the use of corticosteroids, for that small proportion of women for whom preterm delivery is expected, were made in a variety of documents with differing degrees of authority. These include reviews, editorials and textbooks. As early as 1979, a review article in *The Lancet* referenced Liggins and Howie, (1972) and stated that 'Corticosteroids should be given to the mother intramuscularly over 48 h: they reduce the incidence and severity of idiopathic respiratory distress syndrome (IRDS) without substantial risk to mother or fetus' (Ritchie & McClure, 1979, p. 1228). In an editorial statement accompanying the publication of the findings of the NIH's Collaborative Group in the *Journal of Pediatrics* in 1984, Avery's endorsement was particularly strong: she suggested that failure to act on the evidence 'constitutes poor practice' (Avery, 1984, p. 240). Four series of textbooks in the library of the RCOG that had editions from the 1970s through to recent times were examined. Three included some mention of the use of antenatal corticosteroids in the first edition to be published after Liggins and Howie's 1972 paper: in the USA, the third edition of *Danforth's Obstetrics and Gynecology* (Benson, 1977, p. 624); and, in the UK, the second edition of what became *Dewhurst's Textbook of Obstetrics and Gynaecology* (Dewhurst, 1976) and the 13th edition of *Obstetrics by Ten Teachers* (Clayton, Lewis, & Pinker, 1980).

Generally, it was not until the publication in the systematic review of clear conclusions in favour of the use of corticosteroids that more authoritative endorsements in the form of policy recommendations or clinical guidelines were produced. Such statements are important secondary outputs in terms of Buxton and Hanney's framework. In the UK, the use of corticosteroids was recommended in two sets of guidelines produced in 1992 (Joint Working Group of the British Association of Perinatal Medicine and the Research Unit of the Royal College of Physicians, 1992; Royal College of Obstetricians and Gynaecologists (RCOG), (1992). These guidelines were among a small number subsequently included in a policy statement by the National Health Service Management Executive (1993) that took the form of an Executive Letter advocating the greater use of research-based evidence.

In the USA, the 1994 Consensus Conference produced recommendations for use following a year of study and preparation (NIH Consensus Conference, 1995). These recommendations were generally endorsed by the American College of Obstetricians and Gynecologists (Leviton et al., 1999). The conference statement also suggested there could be cost savings of 'more than $3000 per treated neonate' (NIH Consensus Conference, 1995, p. 416).

Assessing and encouraging adoption/application rates

The Buxton and Hanney model suggests there can be a flow of research impacts. This goes from the primary outputs such as publications containing research findings, to the secondary outputs such as clinical policies and guidelines and then into adoption or application by practitioners. The model recognises, however, that the reality is rarely a simple linear sequence. Assessments of use in relation to corticosteroids show a somewhat curious pattern: in at least some countries there was quite a high early adoption of the use of corticosteroids. A survey in 1987 by the Royal Australian College of Obstetricians and Gynaecologists showed that 76% would prescribe antenatal steroids in uncomplicated preterm labour, and a decade later the figure was 97% (Quinlivan, Evans, Dunlop, Beazley, & Newnham, 1998). There are problems when making comparisons between countries, and over time. For example, some surveys of use refer to the percentage of clinicians using it, which in itself can involve self-reporting bias, and others to the percentage of relevant mothers to whom it was administered. And even for those in the latter category there are different interpretations of boundaries for the eligibility criteria. Nevertheless, the analysis below shows that in the UK and USA, even after the evidence for its use had become much firmer, there was considerable resistance to the approach.

Application of the findings in the UK

In 1980 (after Liggins and Howie and only a few other trials) a survey in the UK revealed that as many as 42% of RCOG Members and Fellows claimed to use the treatment frequently and 40% sometimes (Lewis, de Swiet, Boylan, & Bulpitt, 1980). Despite the 51% response rate and possible self-reporting bias, the survey

was seen as the best evidence at the time and preceded any national clinical policies. Figures gathered from several exercises suggest there is little evidence of much, if any, increase in use during the 1980s in the UK (Mugford, 1993) and the low levels of application revealed in some of these exercises stimulated various reactions. The Getting Research into Practice (GRiP) project in Oxford Regional Health authority in the UK attempted to encourage implementation of four pieces of research evidence, of which the use of corticosteroids was one. That project found that implementation at one hospital, the John Radcliffe Hospital in Oxford, had in fact already increased considerably in 1992 (Dopson, Mant, & Hicks, 1994). As part of a growing emphasis on implementation programmes, a study of evidence-based health care was undertaken in three maternity units in one health district of London. It suggested that by the mid 1990s the policy in all three was to give corticosteroids routinely when there was a risk of preterm delivery (Berrow, Humphrey, & Hayward, 1997).

The use of corticosteroids was also included as one of a number of procedures examined in a project conducted in the late 1990s as part of NHS R&D Implementation Methods Programme. This particular study (Wilson et al., 2002) examined changes in levels of compliance with evidence-based recommendations in obstetrics. The study reviewed records to reveal that levels of compliance with the various procedures were very low in 1988; ranging from 0%–23% for RDS in the 20 units studied. The figures rose considerably by 1996, to a median of 82% for RDS. Wilson et al.'s implementation project, in turn, had high media coverage and is already making some impact on midwives (Hanney et al., 2003b). These figures are broadly consistent with a survey in 1997 of 210 obstetric units in the UK that indicated almost all units administered prophylactic antenatal corticosteroids when there was a risk of preterm delivery (Brocklehurst, Gates, McKenzie-McHarg, Alfirevic, & Chamberlain, 1999).

Application of the findings in the USA

The NIH organised a Consensus Conference in 1994 to produce recommendations because of concerns that the rate of adoption was then only 15% (NIH Consensus Conference, 1995). The Agency for Health Care Policy and Research funded an RCT to compare the dissemination of these recommendations in the usual, passive way with active dissemination consisting of a year long education effort led by an influential physician and a nurse co-ordinator at each facility. The results reveal the comparatively low use in the USA even in the mid-1990s; the perhaps surprising success of passive dissemination—adoption rates up from 33% to 58%; and the even greater success of active dissemination—up from 33% to 68% (Leviton et al., 1999).

Reasons for delays in implementation

The reasons for the apparent stalling of the uptake of the use of corticosteroids could be related to various factors, including a critical editorial in the *British Medical Journal* by Roberton (1982). The nature of the debate in the editorial, and the subsequent letters, support the conclusion promulgated from 1992 onwards by the Cochrane Collaboration that the picture would have been clear a decade earlier had a meta-analysis such as that featured in the organisation's logo been published (www.cochrane.org/logo/logoexplanation.htm). Considerable variation between hospitals has been noted, with evidence that it is larger centres at the forefront of medical research that are more likely to have seen an early introduction of corticosteroids (Donaldson, 1992; Leviton et al., 1999).

In an analysis that raised issues that were then addressed by the NIH Conference, Leviton, Baker, Hassol and Goldenberg (1995) described some of the reasons for the low uptake in the USA. They claimed that 'many clinicians may overestimate the probability of negative outcomes resulting from corticosteroid use and underestimate the probability of positive outcomes' (Leviton et al., 1995, p. 315). A key conclusion was that neonatologists were much more supportive of the use of corticosteroids than were obstetricians.

Van Someren (1998) compared the comparatively slow introduction of corticosteroids, which are administered by obstetricians, with the much more rapid introduction of another way of addressing RDS: administration of exogenous surfactant by paediatricians to neonates suffering from RDS. Many factors were thought to account for this, including the fact that the pharmaceutical companies' interest in surfactant also meant the industry had been prepared to fund much larger trials than had occurred with steroids and, therefore, more clinicians had been involved and felt ownership of the trial results.

Assessments of payback

Assessing benefits from Liggins' research

Applying some elements of the multidimensional categorisation and an historical perspective, we start by looking at the benefits from the original work from Liggins, initially funded by a grant of less than £20000 from the Wellcome Trust (Grant et al., 2003). In terms of knowledge production, the original work on sheep resulted in Liggins' (1969) paper that has been cited over 500 times. By itself such basic work is most unlikely to result in applications; instead, it is more likely to inform further research.

This was spectacularly done in this case with Liggins himself making the vital jump from the animal science to the human trial. In this case, therefore, the findings were not only fed into the pool of knowledge and used by many others, but also fed back directly to the work of the original scientist. Indeed, one of the great achievements of Liggins could be said to be that he conducted 'randomisation of the first patient', as advocated by Chalmers (1975). The 1972 paper from Liggins and Howie has been cited over 1200 times, and the analysis from the Wellcome Trust confirms the importance of Liggins' work on subsequent research (Grant et al., 2003).

Somewhat unusually, the 1972 paper began to have an impact not only on subsequent trials, as it proposed should happen, but also on practice. It is possible that the 1980 survey in the UK overstated the degree of use of corticosteroids because, for example, it was likely to be the most research-aware who had formed the majority of the 51% of the Fellows and Members of the RCOG who responded to the questionnaire. Nevertheless, even taking such bias into account, the figures from that survey (42% used it frequently and 40% sometimes) suggest an overall figure that would be in the range of the numbers that should, according to the S-curve, have automatically led to an accelerated adoption rate. It is not entirely clear why there was this disjunction between the initial adoption and the subsequent stalling, if not actual decline, in the 1980s. It is possible that in 1980 some respondents were influenced by the recommendation in the article in *The Lancet* (Ritchie & McClure, 1979), but that after that use trailed off because there was no real follow-up activity until later and there was the negative editorial by Roberton (1982).

Despite the delays, the benefits from Liggins' work were clearly significant in terms of lives saved. A key difference between our payback model and others discussed earlier is that we attempt to explore actual rates of application of research findings and, by addressing issues of what might have caused the uptake, are able to give some indications as to which pieces of research, or implementation activities, played a part in achieving this. Mugford's (1993) analysis estimated the number of neonatal deaths that might be averted at various percentage levels of uptake. It estimated an uptake level at the start of the 1990s of no higher than 20%. Nevertheless, such a figure would translate to an annual averting of over 150 neonatal deaths in England and Wales.

Assessing the benefits from subsequent studies

Identifying the work of Liggins as the starting point is reasonably clear-cut, despite the fact that he was building on a great deal of earlier research from others. Within the assessment of the impact from the whole stream of research that started with the 1969 paper, it is more difficult to estimate the benefits from specific projects that followed the work of Liggins and colleagues. The NIH-funded trial was second only to Liggins and Howie in terms of the numbers included in their study and in this sense it was significant. A 1986 NIH assessment of the benefits did so in terms of cost savings. Based on the NIH-funded trial it claimed, somewhat unfairly, that the National Heart, Lung and Blood Institute 'developed antenatal steroid therapy to prevent neonatal respiratory distress syndrome' (NIH, 1993). The cost of the research between 1976 and 1983 was $7.4 million (in 1992 figures) and, with the analysis partly based on speculation about take-up rates, the potential reduction in treatment costs was claimed to be between $16.5 million and $145.0 million. This analysis can be criticised on methodological grounds because no attempt was made to assess actual levels of implementation—either before the findings of the NIH study were produced, or sometime afterwards—but it did usefully highlight the potential benefits from full implementation.

The next major development on which payback analysis could focus is the systematic review. The original review article has been cited on about 370 occasions. In terms of impact, the major expansion in the use of corticosteroids occurred after the systematic review. The impact, however, took some time and, at least in the UK and USA, also involved further activities by parts of the health research system.

In this context it seemed appropriate to Chalmers and Mugford that Mugford should produce a brief paper for the Department of Health based on her economic evaluation (Mugford et al., 1991). This appears to have been used in various official documents, a key one being from the R&D Division of the Department of Health (Department of Health, 1993). This document, like the whole new NHS R&D strategy, stressed the importance of the utilisation of research findings, but the use of corticosteroids was the only research application for which it gave a calculation. It described the systematic review but claimed the then take-up rate was only 15–20%. Based on the calculations of the proportion of lives saved by use of corticosteroids, and figures about cost savings, it stated that if the information 'on antenatal steroids in premature babies were to be fully utilised, it is estimated that resources equivalent to up to 5% of current spending on neonatal hospital care might be made available for use elsewhere in that field of health care' (Department of Health, 1993, p. 35).

The benefits from the economic evaluation by Mugford et al. (1991) were assessed as one of Buxton and Hanney's payback case studies (Hanney, 1994). The cost-effectiveness study itself cost very little, but was an advance because none of the antenatal steroid trials had

included formal economic evaluation even though some reported hospital charges. Applying the payback analysis just to Mugford's cost-effectiveness study highlights the difficulty of attempting to separate out the benefits of one item from the wider stream of work. In general, the documentary review and interviews with key informants established that the systematic review was the evidence that was having the most impact on the production of secondary outputs in the form of decisions by professional groups and administrators to produce their guidelines. Furthermore, the analysis in the payback study demonstrated the difficulties of realising the cost savings in contracts: such savings would mostly be absorbed by providing better or alternative care to other babies (Hanney, 1994).

Guidelines and additional studies such as GRiP will have contributed to increasing usage throughout the 1990s but, as noted by Wilson et al. (2002), it is difficult to identify precise factors leading to the increased uptake. Mugford's 1993 analysis indicated that if uptake in England and Wales rose from a possible 20% figure to 75% of eligible cases, this could result in over 400 additional deaths being averted annually. Data described earlier suggest that adoption rates are now above 75%, therefore there has clearly been substantial benefit in terms of lives saved and release of resources for other uses. Analysis by Mugford et al. (1991) was also used by Wilson et al., but related to the whole UK, to suggest that even the 1996 uptake figures of about 80% still implied 'approximately 500 avoidable cases of RDS and 200 avoidable deaths from prematurity each year in the UK' (Wilson et al., 2002, p.180). Evidence from the way in which uptake increased in the 1990s implies further movement towards 100% adoption rate and, therefore, even more neonatal deaths averted each year. There is, however, a complication. The much greater use of surfactant in the 1990s, and improved neonatal intensive care services, mean that even without corticosteroids many of the deaths that would previously have occurred might now be averted. This makes it difficult to state that without corticosteroids there would be a specific number of neonatal deaths. Nevertheless, not all the mortality and morbidity would be avoided by relying on surfactant and the NIH Consensus Conference stated: 'The benefits of antenatal corticosteroids are additive to those derived from surfactant therapy' (1995, p. 417). Furthermore, the use of corticosteroids as a preventative therapy is still recommended in the most recent editions of textbooks (Edmonds, 1999; Scott, Gibbs, Karlan, & Haney, 2003).

Discussion and conclusion

This analysis shows that the body of research examined led to important knowledge production, or primary outputs, and resulted in a string of significant secondary outputs in the form of guidelines. But above all, and key to providing 'good stories' to assist with justifying research funds, this example provides an account of how health research can lead to health gains in terms of reduced mortality and morbidity.

The analysis presented here has limitations. It is based on a single case study of a topic that is exceptional in various ways including: a rapid move from animal to human studies; an involvement in the development of the Cochrane Collaboration; a gain both in terms of health and cost reduction; and many existing studies on which to draw. Furthermore, the type of issue involved, i.e. a clinical intervention that can be studied by traditional RCTs, is one where this type of payback analysis should be most feasible, and yet even here it proved very difficult to give precise figures about the level of payback achieved. Nevertheless, we would argue that the very difficulties involved show how in practice a detailed analysis of actual adoption rates and the factors behind them is likely to be necessary for a realistic analysis of the payback achieved from the various elements in any line of research and its implementation.

In terms of how to conduct assessments, this analysis shows it is sometimes possible to follow through a series of stages from basic to clinical science and on to some implementation in clinical practice. As a case study it also demonstrates that a multidimensional approach to assessing benefits has advantages, in terms of its flexibility, in dealing with what can become complex situations. Limiting the analysis to one category of benefits reduces the scope of the analysis as when, for example, it becomes difficult to demonstrate features such as actual cost savings. By considering a series of stages, the assessment of benefits can also help illustrate points in an holistic way that complements, and builds upon, some of the more detailed analyses undertaken by other researchers of specific aspects of the utilisation processes. It allows the more diffuse and very important contextual issues to be considered. This case suggests that there can be periods when the implementation appears to stall rather than automatically accelerate as would be indicated by the S-curve. This illustrates that those running health research systems need carefully to consider how they can best contribute to encouraging implementation.

The material described here could help to build a counterfactual scenario. It is possible to imagine that without follow-on activities after the RCTs there would have been a moderate amount of implementation of the research findings, but less than the desirable levels now being achieved. The role of the systematic review was shown to be particularly important in moving application rates to higher levels. The timing of this therapy's introduction complicates the analysis as it straddles the development of meta-analytic approaches. Nevertheless,

this study usefully illustrates the crucial importance of such methods but at the same time highlights, as noted repeatedly by those preparing such reviews, that they are not sufficient. Some appropriate implementation is always likely to occur prior to systematic reviews, and there is still resistance to implementation after systematic reviews. Other approaches are also required.

Economic evaluations have various roles to play. Given the rather unusual finding that this therapy offered both potential cost savings as well as health gains, economic evaluations therefore played the role of assisting promotion of the introduction of corticosteroids. The production of guidelines, policies, consensus conference statements and active implementation strategies also all played a part and illustrate the frequent need for multiple approaches towards research implementation (Haines & Jones, 1994). Indeed, our case study suggests some of these approaches can be more successful in certain circumstances than is often thought. This, in turn, reinforces the need for a multidimensional approach to the *assessment* of benefits over a long timescale.

Finally, therefore, this case study shows how assessments of the benefits from health research can fulfil the roles described in the Introduction, including contributing towards greater understanding of the processes of research utilisation. An advantage of payback assessment approaches, compared with some other utilisation studies, is that they can, as here, provide a rather more positive message in that their focus is primarily on what has been achieved, even if full utilisation is still awaited. Hence, their potential usefulness in providing justification for current levels of expenditure on health research and perhaps in advocacy for extended funding. In this case we have shown that an often quoted example of underutilisation of research findings can, nevertheless, be used to demonstrate the considerable benefits that can arise from various aspects of a stream of research.

Acknowledgements

The authors gratefully acknowledge the encouragement and helpful comments on earlier drafts of this paper received from Iain Chalmers, Ross Howie, Shyama Kuruvilla, Graham Liggins, Bryony Soper, two referees, and participants at the Scientific Basis of Health Services Conference, 2003, where some of these ideas were presented. We are also grateful for the contribution to the Wellcome Trust's Comroe and Dripps study made by Barbara Mason and Liz Green and for funding from the Department of Health's Policy Research Programme and the NHS Executive's London Office. This funding supported parts of the analysis undertaken by the authors.

References

Antman, E. M., Lau, J., Kupelnick, B., Mosteller, F., & Chalmers, T. C. (1992). A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *Journal of the American Medical Association*, *268*(2), 240–248.

Avery, M. E. (1975). Pharmacological approaches to the acceleration of fetal lung maturation. *British Medical Bulletin*, *31*(1), 13–17.

Avery, M. E. (1984). The argument for prenatal administration of dexamethasone to prevent respiratory distress syndrome. *Journal of Pediatrics*, *104*(2), 240.

Benson, R. (1977). Premature labor and the small-for-dates infant. In D.N. Danforth (Ed.), *Obstetrics and gynecology*, (3rd ed.). (pp. 618–627). Hagerstown MD:Harper & Row.

Berrow, D., Humphrey, C., & Hayward, J. (1997). Understanding the relation between research and clinical policy: A study of clinicians' views. *Quality in Health Care*, *6*, 181–186.

Brocklehurst, P., Gates, S., McKenzie-McHarg, K., Alfirevic, Z., & Chamberlain, G. (1999). Are we prescribing multiple courses of antenatal corticosteroids? A survey of practice in the UK. *British Journal of Obstetrics and Gynaecology*, *106*, 977–979.

Buxton, M., & Hanney, S. (1996). How can payback from health services research be assessed? *Journal of Health Services Research and Policy*, *1*(1), 35–43.

Chalmers, T. C. (1975). Randomization of the first patient. *Medical Clinics of North America*, *59*(4), 1035–1038.

Clayton, S. G., Lewis, T. L. T., & Pinker, G. (Eds.), (1980). *Obstetrics by ten teachers* (13th ed.). London: Edward Arnold.

Cochrane Collaboration. (2003). Brochure. [http://www.cochrane.org/cochrane/cc-broch.htm –accessed 29 May 2003].

Collaborative Group on Antenatal Steroid Therapy. (1981). Effect of antenatal dexamethasone administration on the prevention of respiratory distress syndrome. *American Journal of Obstetrics and Gynecology*, *141*, 276–287.

Collaborative Group on Antenatal Steroid Therapy. (1984). Effect of antenatal dexamethasone administration on the infant: Long-term follow-up. *Journal of Pediatrics*, *104*(2), 259–267.

Comroe, J. H., & Dripps, R. D. (1976). Scientific basis for the support of biomedical science. *Science*, *192*, 105–111.

Crowley, P. (1989). Promoting pulmonary maturity. In I. Chalmers, M. Enkin, & M. J. N. C. Keirse (Eds.). *Effective care in pregnancy and childbirth* (pp. 746–764). Oxford: Oxford University Press.

Crowley, P. (1995). Antenatal corticosteroid therapy: A meta-analysis of the randomized trials, 1972 to 1994. *American Journal of Obstetrics and Gynecology*, *173*(1), 322–334.

Crowley, P. (2002). Prophylactic corticosteroids for preterm birth. (Cochrane review) In *The Cochrane Library, Issue 4*. Oxford: Update Software.

Crowley, P., Chalmers, I., & Keirse, M. J. N. C. (1990). The effects of corticosteroid administration before preterm delivery: An overview of the evidence from controlled trials. *British Journal of Obstetrics and Gynaecology*, *97*, 11–25.

Croxson, B., Hanney, S., & Buxton, M. (2001). Routine monitoring of performance: What makes health research and development different? *Journal of Health Services Research and Policy*, *6*(4), 226–232.

Department of Health. (1993). *Research for health*. London: Department of Health.

Dewhurst, C. J. (Ed.). (1976). *Integrated obstetrics and gynaecology for postgraduates*. (2nd ed.). Oxford: Blackwell Scientific Publications.

Donaldson, L. J. (1992). Maintaining excellence: The preservation and development of specialised services. *British Medical Journal*, *305*, 1280–1283.

Dopson, S., Mant, J., & Hicks, N. (1994). Getting research into practice: Facing the issues. *Journal of Management in Medicine*, *8*(6), 4–12.

Drummond, M. F., Davies, L. M., & Ferris, F. L. (1992). Assessing the costs and benefits of medical research: The diabetic retinopathy study. *Social Science & Medicine*, *34*(9), 973–981.

Edmonds, K. (Ed.). (1999). *Dewhurst's textbook of obstetrics and gynaecology for postgraduates*. (6th ed.). Oxford: Blackwell Science.

Graham, I. (1997). *Episiotomy: Challenging obstetric interventions*. Oxford: Blackwell Science.

Grant, J., Cottrell, R., Cluzeau, F., & Fawcett, G. (2000). Evaluating "payback" on biomedical research from papers cited in clinical guidelines: Applied bibliometric study. *British Medical Journal*, *320*, 1107–1111.

Grant, J., Green, L., & Mason, B. (2003). *From bedside to bench: Comroe and Dripps revisited. HERG Research Report No. 30*. Uxbridge: Brunel University.

Grimshaw, J. M., Shirran, L., Thomas, R., Mowatt, G., Fraser, C., Bero, L., Grilli, R., Harvey, E., Oxman, A., & O'Brien, M. A. (2001). Changing provider behaviour: An overview of systematic reviews of interventions. *Medical Care*, *39*(Suppl. 2), 12–45.

Haines, A., & Jones, R. (1994). Implementing findings of research. *British Medical Journal*, *308*, 1488–1492.

Hanney, S. (1994). Economic evaluation of the prevention of neonatal respiratory distress syndrome. In M. Buxton, & S. Hanney (Eds.), *Assessing payback from Department of Health Research & Development: Eight case studies. Research Report No. 19*, Vol. 2. (pp. 39–48). Uxbridge: HERG, Brunel University.

Hanney, S., Gonzalez-Block, M., Buxton, M., & Kogan, M. (2003a). The utilisation of health research in policy-making: Concepts, examples and methods of assessment. *Health Research Policy and Systems*, 1 (2) [http://www.health-policy-systems.com/content/1/1/2].

Hanney, S., Packwood, T., & Buxton, M. (2000). Evaluating the benefits from health research and development centres: A categorization, a model and examples of application. *Evaluation: The International Journal of Theory, Research and Practice*, *6*(2), 137–160.

Hanney, S., Soper, B., & Buxton, M. (2003b). *Evaluation of the NHS R&D Implementation Methods Programme. HERG Research Report No. 29*. Uxbridge: HERG, Brunel University.

Haynes, B., & Haines, A. (1998). Barriers and bridges to evidence based clinical practice. *British Medical Journal*, *317*, 273–276.

Joint Working Group of the British Association of Perinatal Medicine and the Research Unit of the Royal College of Physicians. (1992). Development of audit measures and guidelines for good practice in the management of neonatal respiratory distress syndrome. *Archives of Disease in Childhood*, *67*, 1221–1227.

Kogan, M., & Henkel, M. (1983). *Government and research: The Rothschild experiment in a government department*. London: Heinemann.

Leviton, L. C., Baker, S., Hassol, A., & Goldenberg, R. L. (1995). An exploration of opinion and practice patterns affecting low use of antenatal corticosteroids. *American Journal of Obstetrics and Gynecology*, *173*, 312–316.

Leviton, L. C., Goldenberg, R. L., Baker, C. S., Schwartz, R. M., Freda, M. C., Fish, L. J., Cliver, S. P., Rouse, D. J., Chazotte, C., Merkatz, I. R., & Raczynski, J. M. (1999). Methods to encourage the use of antenatal corticosteroid therapy for fetal maturation. A randomized controlled trial. *Journal of American Medical Association*, *281*(1), 46–52.

Lewis, P. J., de Swiet, M., Boylan, P., & Bulpitt, C. J. (1980). How obstetricians in the United Kingdom manage preterm labour. *British Journal of Obstetrics and Gynaecology*, *87*, 574–577.

Liggins, G. C. (1969). Premature delivery of foetal lambs infused with glucocorticoids. *Journal of Endocrinology*, *45*, 515–523.

Liggins, G. C., & Howie, R. N. (1972). A controlled trial of antepartum glucocorticoid treatment for prevention of the respiratory distress syndrome in premature infants. *Pediatrics*, *50*(4), 515–525.

Little, B. (1981). Editorial comment. *American Journal of Obstetrics and Gynecology*, *141*, 286.

Mugford, M. (1993). Corticosteroids in preterm delivery—estimated impact on neonatal outcomes and health service resources. In Oxford Regional Health Authority, *A review of stillbirths and neonatal mortality in the Oxford Region, 1992* (pp.37–40). Oxford: Oxford Regional Health Authority.

Mugford, M., Piercy, J., & Chalmers, I. (1991). Cost implications of different approaches to the prevention of respiratory distress syndrome. *Archives of Disease in Childhood*, *66*, 757–764.

National Health Service Management Executive. (1993). Improving Clinical Effectiveness. EL (93)115. Leeds: NHS.

National Institutes of Health. (1993). *Cost savings resulting from NIH research support* (2nd ed). Bethesda, MD: NIH.

National Institutes of Health Consensus Conference. (1995). Effect of corticosteroids for fetal maturation on perinatal outcomes. *Journal of the American Medical Association*, *273*(5), 413–418.

Quinlivan, J. A., Evans, S. F., Dunlop, S. A., Beazley, L. D., & Newnham, J. P. (1998). Use of corticosteroids by Australian obstetricians—a survey of clinical practice. *The Australian and New Zealand Journal of Obstetrics and Gynaecology*, *38*(1), 1–7.

Ritchie, K., & McClure, G. (1979). Prematurity. *The Lancet*, *2*, 1227–1229.

Roberton, N. R. C. (1982). Advances in respiratory distress syndrome. *British Medical Journal*, *284*, 917–918.

Rogers, E. (1995). *Diffusion of innovations* (4th ed.). New York: The Free Press.

Royal College of Obstetricians and Gynaecologists. (1992). *Antenatal corticosteroid administration reduces the incidence of neonatal respiratory distress syndrome*. London: Scientific Advisory Committee, Royal College of Obstetricians and Gynaecologists.

Scott, J. R., Gibbs, R. S., Karlan, B. Y., & Haney, A. F. (Eds.). (2003). *Danforth's obstetrics and gynecology*. (9th ed.). Philadelphia: Lippincott Williams & Wilkins.

Sculpher, M., Drummond, M., & Buxton, M. (1997). The iterative use of economic evaluation as part of the process of health technology assessment. *Journal of Health Services Research & Policy*, *2*, 26–30.

Smith, R. (1987). Comroe and Dripps revisited. *British Medical Journal*, *295*, 1404.

Smith, R. (2001). Measuring the social impact of research. *British Medical Journal*, *323*, 528.

Townsend, J., Buxton, M., & Harper, G. (2003). Prioritisation of health technology assessment. The PATHS model: Methods and case studies. *Health Technology Assessment*, *7*(20).

Van Someren, V. (1998). Changing clinical practice in the light of the evidence: Two contrasting stories from perinatology. In A. Haines, & A. Donald (Eds.). *Getting research findings into practice* (pp. 143–151). London: BMJ Books.

Wilson, B., Thornton, J. G., Hewison, J., Lilford, R. J., Watt, I., Braunholtz, D., & Robinson, M. (2002). The Leeds university maternity audit project. *International Journal for Quality in Health Care*, *14*(3), 175–181.