SOCIOLOGY OF HEALTH & ILLNESS

Sociology of Health & Illness Vol. 29 No. 2 2007 ISSN 0141–9889, pp. 180–197
doi: 10.1111/j.1467-9566.2007.00531.x

Entangled evidence: knowledge making in systematic reviews in healthcare

Tiago Moreira

School of Applied Social Sciences, Durham University

Abstract As the volume of biomedical information escalates and its uses diversify, systematic reviews and meta-analyses – the compilation, selection and statistical analysis of pooled results from similar studies – are becoming an increasingly accepted method in the evaluation of healthcare technologies and interventions. We thus observe a proliferation of laboratories conducting this type of research. How is knowledge constructed in systematic reviews and meta-analysis in healthcare? Drawing on ethnographic data collected during 18 months of fieldwork in a research centre devoted to the development of evidence-based clinical-practice guidelines and systematic reviews, the paper argues that knowledge construction in secondary research in healthcare is structured upon a parallel process of disentanglement and qualification of data. In disentanglement, knowledge practices attempt to extricate data from the milieus in which they are commonly found (databases, texts, other research centres, etc.). In qualification, the focus of activities is on endowing data with new qualities – such as precision, unbiasness or 'fairness' – through the use of templates, graphical platforms and techno-political debates. The accomplishment of these two processes is fundamental to establishing the persuasive power that meta-analyses appear to have in contemporary healthcare politics.

Keywords: systematic reviews, knowledge practices, ethnography, healthcare

Introduction

At the end of the 20th century, systematic reviews – the compilation, selection and analysis of pooled results from similar studies – and meta-analysis – the statistical calculation of those pooled results – became progressively central to the processes of debate and decision making in healthcare (Mulrow 1994). This was accompanied both by a proliferation of laboratories conducting systematic reviews (Levin 2001), and methodological research on the techniques and procedures of secondary research (Egger, Davey Smith and Altman 2001). Little is known, however, about how, in practice, knowledge is composed in these laboratories or about how this knowledge comes to mediate the relationship between health research and health politics in contemporary democracies. Drawing on ethnographic data and the anthropology of calculation (Callon 1998a), I argue, in this paper, that knowledge making in systematic reviewing is structured upon parallel attempts a) to *disentangle* data from the milieus in which they are commonly found (databases, texts, other research

Published by Blackwell Publishing Ltd., 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

centres), and b) to re-*qualify* that data through comparisons across a variety of 'platforms' (tables, graphs, equations, and political controversies). I suggest that the accomplishment of these two processes is fundamental in establishing the persuasive power of systematic reviews and *equipping* political action in contemporary health care.

In the first section of the paper I describe how the link between the political organisation of healthcare and the practices of 'evidence' production has historically come to rely on the formal methods of evaluation of health technology assessment, and amongst these, increasingly on systematic reviewing and meta-analyses. Then, after a brief methodological note, I explore what I identify as 'the practical problem' of systematic reviewing in the way in which reviewers frame their interaction with their main research material – texts. In the main section of the paper I explore the procedures used by reviewers in order to make these texts 'docile', and how this is intimately connected with the graphic and organisational apparatus through which reviewers compose their comparisons and recalculations. I then move on to describe how systematic reviews are used in the development of evidence-based clinical guidelines. In this section, I explore how the trajectory of the systematic review depends on a skilled understanding of the techno-political debates that surround healthcare interventions. I argue that the embedding of the systematic review in this techno-political platform in turn enables systematic reviews to 'frame' the character of those debates and provides a key insight into how reviews come to structure and sustain the interaction between science and politics in healthcare.

Systematic reviews and the politics of healthcare

The consolidation of systematic reviewing as a research field can be seen as the outcome of, but also a key factor in, important changes in the organisation of health-care. First, socio-demographic change initiated during the 18th and 19th centuries became crystallised in growing sectors of the population of Western societies suffering from lasting, expensive chronic illnesses. Then, in the post war period, medicine was seen as having shifted from producing dramatic changes in the lives of humans to introducing small incremental variations in the patterns of the health of populations (Mckeown 1979). This was concurrent with rising healthcare costs (Abel-Smith 1996). In the 1980s, healthcare delivery became increasingly buyer driven – either by insurance companies, trusts or the patients themselves – and its deployment progressively taken away from direct providers, 'rationalised' (Light 2000) and mediated by economic calculations (Ashmore, Mulkay and Pinch 1990). In addition, the position of the medical profession in relation to its constituencies changed in that its regulation within healthcare systems in western societies was progressively delegated (Harrison 2001, 1999, Light and Levine 1988).

In this period, medical knowledge has also become a public issue in the controversy over the value of evidence-based medicine (EBM). The problems and uncertainties of medical knowledge and its boundaries were addressed, particularly in the US, by an increasing number of practice guidelines produced by speciality societies combining 'overviews of the literature' with more traditional 'expert opinions' (Timmermans and Berg 2003). In the UK, the EBM debate was central to the political evolution of the healthcare system from the middle of the 1990s (Ghali and Sargious 2002, Harrison 1998) leading to the establishment of the National Institute for Clinical Excellence, a Special Health Authority within the National Health Service that provides primary and acute care trusts, patients and clinicians with guidance – in the form of health technology assessments and clinical guidelines – on 'best practice' (Rawlins 1999). It is important to remember though that the changes in the

organisation of biomedical knowledge associated with EBM emerged at the tail end of a historical trend of 'mathematization' of medical knowledge (Marks 1999, Porter 1995). The use of statistics was particularly successful in medicine by differentiating spaces, populations, diseases and interventions, and by enabling comparison between those. Among these techniques the randomised controlled trial became a particularly robust statistical experimental method. By deploying a public, fair comparison between interventions, this research method influenced the organisation of legitimate decision making in post-war liberal democracies (Marks 1997). They were thus encouraged by the State through legislation on pharmaceuticals and consistent funding policies of academic medicine and medical education. In this process, the alliance between modern political organisation of healthcare and the techniques of statistical comparison became increasingly strong (Lehoux and Blume 2000).

The interest in research-synthesis techniques in the 1990s is motivated explicitly by these changes in the politico-epistemic requirements of precision and fairness of studies of health interventions. Although the techniques of pooling and comparison of experimental results from different studies had been used since the 19th century both in medicine and in other disciplines (Chalmers, Hedges and Cooper 2002), it is only during the 1970s and 1980s that they come to be developed as a routine method of research. This development was specifically located in the field of applied social science and social policy as practised in the US Program Evaluation and Methodology of the General Accounting Office (Hunt 1997). Its transposition to contemporary medicine is marked by the study of streptokinase, a clot-busting enzyme to treat acute myocardial infarction (AMI) (Yusuf *et al.* 1985). This cumulative meta-analysis of AMI therapies led to the claim that the statistical analysis of a relatively small number of minor studies could show results similar to the results of a large-scale study. More recently, however, the justification for the use of meta-analyses in medicine has been increasingly associated with its power to summarise the enormous amount of information that is produced about health interventions and to evaluate the significance of possibly conflicting sets of information.

Building upon the configuration of factors described above, systematic-review methodologists and practitioners appear to have been able to persuade various constituencies about the value of their products to support healthcare and policy decisions. In this, systematic reviewing can be seen to partially stabilise, by technical means, the boundary and interactions between health politics and medical science (Tanenbaum 1994). This boundary role might help explain the sustained growth in use of meta-analysis in medicine in the last two decades (see Figure 1).

This increase has been concurrent with the proliferation of centres conducting trials, systematic reviews and meta-analyses and the establishment of professional networks of communication and methodological training, the most influential of which has been the Cochrane Collaboration. This proliferation has, however, been a contested process, done against a constant backdrop of criticisms focusing on the parasitic character of secondary research. Such criticisms have in turn led to methodological research concentrating on the 'added value' that systematic reviews can deliver. Most published research about systematic reviews and meta-analyses has thus been focused on the development and appraisal of methodological procedures and statistical calculations. It has also mainly been conducted by practitioners for practitioners, in order to consolidate the methodological robustness of these practices. This research informs, influences and *prescribes* the ways in which systematic reviewing should be conducted and reported.

This paper is dedicated instead to the *description* of the knowledge practices of systematic reviewing. It identifies the epistemic challenges and solutions of reviewing research from the

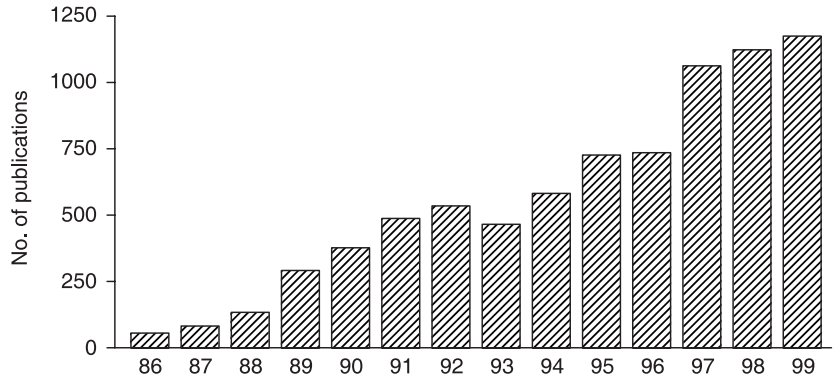

Figure 1 Number of publications concerning meta-analysis, 1986–1999. Results from MEDLINE search using text word and medical subject (MESH) heading 'meta-analysis' and text word 'systematic review'.
Source (Egger, Davey Smith and O'Rourke 2001), used with kind permission of authors.

point of view of everyday practice. This empirical approach is encouraged by practitioners' reports on the dilemmas and challenges of systematic reviewing. In a *BMJ* paper entitled 'The private life of systematic reviews', for example, Roberts and Schierhout say:

> In contrast to the public nature of the more technical aspects of a systematic review, details of these interpersonal interchanges [between reviewers and original authors] are never mentioned in the published report. We argue that a closer look at these interpersonal interactions raises important questions for the evolving science of reviewing the medical literature. […]Eliciting unpublished information from trialists is the social challenge of a systematic review (1997: 686).

My suggestion, based on ethnographic data, is that the social challenge identified by Roberts and Schierhout in the particularities of eliciting data from researchers is in fact part of a broader fundamental epistemic challenge faced by systematic reviewers: how to disentangle data from the contexts in which they were produced and/or are stored? This challenge encompasses the whole variety of activities conducted by systematic reviewers:

- the selection and retrieval of relevant data-rich publications from large databases;
- the selection and retrieval of data from selected publication;
- the attempts to obtain unpublished data from trial managers,
- the attempts to find and adequately format the figures required by statistical equations.

In responding to this challenge, reviewers devise and deploy strategies of *disentanglement*. These strategies attempt to break the ties between data and the original milieus where they were produced. Attempts at dissociation are accompanied simultaneously by practices of *qualification*. It is a central concern in systematic reviewing that the results produced do not merely reproduce the research examined. Thus, another of reviewers' main epistemic challenges has to do with the transformation of the data being selected, elicited and abstracted. In this, I am suggesting that reviewers *modify* the data they are working on or with by submitting it to a series of procedures of comparison – between quality markers and methods reported in trial publications, between different trials, between their own

results and those of previous meta-analyses, and between positions in techno-political controversies about the intervention. In this process of adjustment, reviewers are able to re-qualify the data used for statistical calculations in their meta-analyses, transforming the technical and political meaning of the data being analysed. Such a description of systematic reviewing is thus invested in understanding the epistemic dynamic upon which 'evidence' is deconstructed and re-calculated.

To formulate this analytical framework, I draw on Callon's research on economics and its role in enacting the economic calculative agent (Callon 1998b). For Callon, the definition of the agents, goods and platforms of economic activity – its 'framing' – is dependent upon a continuous work of 'cleansing, of disconnection' between these entities and their social and technical contexts (Callon 1998a). In this process of disentanglement, economics plays a fundamental part through its attempts to define what is 'internal' and what is 'external' to the market. Callon's model supports the suggestion that the apparently straightforward work of data abstraction and recalculation, not only changes the political meaning of biomedical data itself but is also instrumental in enabling a particular form of political decision making – sometimes called 'evidence-based politics'. This particular form of politics can be said to be dependent upon these knowledge practices. In this paper, I start investigating this connection by focusing on how the knowledge practices of systematic reviewing frame and become entangled with this form of politics.

My approach is ethnographic: I am interested in the mundane activities through which systematic reviewers are able, at the same time, to 'mine' texts and recalculate data contained therein and to shape these data to fit the logics of ongoing health research and political debates. I follow the texts, objects, data and correlated practices of systematic review from the moment they enter the 'laboratory' through to their difficult circulation in more political arenas such as clinical practice guideline meetings. An important issue in this study had to do with the assumptions brought by an ethnographic approach about the nature of scientific knowledge. Combining both a sociological theory of science and a method for its analysis, the ethnography of science has represented scientific knowledge as a construction effected by mundane, ordinary activities (Knorr-Cetina 1981, Latour and Woolgar 1986, Lynch 1985a). In this, some commentators have suggested that early 'laboratory studies' focused too narrowly in the activities conducted within laboratories, neglecting the translations, negotiations and contestations that scientific facts and the identity of 'science' undergo continuously in diverse public arenas (Callon, Lascoumes and Barthe 2001, Gieryn 1999, Irwin and Michael 2003). What this study offers is an ethnographic description of a set of practices mainly concerned with re-distributing value to scientific data that are already in the public domain. Systematic reviewing is, in this sense, about taking *back to the laboratory* data that have already been produced by other centres of calculation (Latour 1990). This means that I have used the methods of the ethnography of science to suggest that there are changes in the process by which biomedical science is evaluated in public, specifically in the way in which legitimate forms of interaction between science and politics are built in contemporary democracies (Guston 2001).

Setting and methodology

Between 2002 and 2003, I conducted an ethnographic study of a British research unit dedicated to the development of systematic reviews of healthcare. The unit was staffed by the director of the unit (a health economist), a statistician, four systematic reviewers, one information scientist and two administrative assistants. I used a variety of ethnographic

techniques of data collection-analysis: extensive, detailed fieldnotes produced during long periods of immersion in the day-to-day life of the unit; printed materials gathered during team meetings (minutes, agendas, etc.) and training sessions (research papers, statistical formulae, etc.) in which I participated as a member; work-logs written by all team members over the period of one week; in-depth interviews with the research staff of the unit; and tape-recordings of 'reconciliation sessions' between two reviewers as well as access to the records of those sessions kept by the reviewers themselves. The log study involved asking members to keep a descriptive record of their daily activities over the course of two hours per day during five consecutive working days. Interviews were semi-structured, the schedule partially directing the interviewee to reflect upon the information provided in their respective work-log, and their role in the research developed by the unit.

In parallel, I observed a series of clinical practice guideline development meetings for which the members of the unit developed some of their systematic reviews. The guidelines observed were developed for clinical conditions that are largely managed in primary care, though they also addressed areas of clinical practice in secondary care. The meetings gathered a multidisciplinary group of primary- and secondary-care clinicians, pharmacists and patient representatives. The process was explicitly evidence-based, with studies tabulated and summarised by the research team. These were discussed by the groups and then statements of the evidence and accompanying recommendations for practice were formulated and subsequently revised as appropriate. Stringent procedures were necessary to ensure the anonymity of participants in both studies.

The data analysis and collection were strongly interwoven during the year-and-a-half of fieldwork. The overall approach can be best summarised by Jack Katz's notion of analytic induction, where researchers 'search simultaneously for an explanation that will fit all the evidence and for a definition of the problem that, without "cooking" or hiding data, makes relevant only the evidence that fits the explanation' (Katz 2002: 485). The internal validity of the study was accomplished through a 'theoretical saturation' of analytical findings (Strauss and Corbin 1998). Further analytical consistency was achieved by obtaining respondent validation at various stages of the data analysis and of the overall interpretation presented in this paper.

(Back) in the textual laboratory

> The systematic reviewers' office, located on the upper floor of the building, is a relatively small room: three workstations, with two tables each; three PCs; a shared printer; shelves covering two of the main walls; a disused kettle and sink next to the window; and papers, research papers everywhere. On tables, piled in variously sized heaps, filed, binned, underlined. There are papers in communal areas, such as the table next to the printer. These are, a label tells us, 'to be read'. There are papers sorted by types of intervention in different cardboard files on the shelves. Others have already made their way to the metal filing cabinets in the secretary's office. In the corner, lie discarded papers, rejected. Some of them have already been binned. Others have just come in from the interlibrary loan, and are still in their plastic folders. On each of the reviewers' tables lie their own sets of papers, some in the process of being read, others waiting to be read. Some of the former are heavily marked with ink and pencil notes, underlines, highlights, circles, interrogations, comments. Some of them have post-its attached, with more comments. Other papers, closer to the computer, are half marked, with comments such as 'check patient numbers are in ((the other)) paper'. The computer hums in the midst of more

> post-its – 'get Asanti, 1995', 'lost to follow-up!' – attached to the monitor. On the screen a table filled with numbers, percentages, risk ratios.[…]
>
> [Fieldnote #15 Extract – 07/02]

This fieldnote extract captures, first and foremost, my amazement at the volume of paper that is handled, read, marked, classified, managed, stored and moved around as a matter of course every day in a systematic reviewer's office. Being a researcher myself, I am used to paper. The shelves in my office are themselves stacked with files full of research paper of one kind or another. On my desk, as I write, I have various papers and books lying around, some for this paper, others still for the report that I finished last week. There is however a difference between my office and the systematic reviewers' office; that difference is a matter of *volume* that denotes whole different practices of text use. This begs the question: how do texts figure in the social organisation of knowledge in systematic reviewing? Another note, written later in the fieldwork, begins to clarify this question:

> [Theresa, one of the systematic reviewers, told me that] in the review she was doing at that moment the original 'search string' had brought them about 8,000 papers reporting on trials for [chronic illness]. 'Sifting' through these, all these, abstracts was, from her point of view, only possible because they had set selection criteria. She only needed to *look for* specific things, such as the number of participants in the trial, or whether or not the outcomes of the trial were reported ('sometimes it's just a trial protocol'). However, in the end, she still had a larger proportion of 'unsures' than 'inclusions' or 'rejections'. She saw this as a consequence of how authors 'conceal' the real nature of the research, sometimes 'presenting something as a trial when *in fact* it isn't'. She ordered the papers in and went through them one more time, now using more stringent criteria, checking and noting down if they reported on, for example, blinding or concealment of allocation in trial design, or provided information measures of representativity of the trial (*e.g.* age mean and range, percentages of male and female patients). This was the basis for further selection. She finished with a reduced number of 'quality papers' that she was prepared to 'mine' for data.
>
> [Fieldnote #177 Extract – 02/03]

Theresa's approach to texts seems to be underpinned by a strategy of avoidance, or as one might put it, by trying 'not to read' papers. Presented with an immense quantity of possible trial data to analyse, Theresa uses a variety of filters to distinguish 'data rich' papers from non-relevant material such as 'just a trial protocol'. She is also encouraged, in her selection, by her view that authors attempt either to conceal the real nature of the research or to make it 'look better' than it is. In order to get to the data that she is interested in she must disregard these attempts to 'lure' the reviewer into requesting the paper or believing their interpretation of the trial results. The aim of the systematic reviewer is, as another reviewer put it, to be able to 'data abstract a paper and *not know what it is saying*, which treatment [it claims] is better' [Colin, Unit Meeting, 05.03; my emphasis]. However, most of the information that Theresa wants is embedded in the textual attempts to conceal or embellish the results of the trial. In this lies the *practical problem* that is at the core of systematic reviewing: *how to resist the enticements, persuasive strategies and the arguments inscribed in papers while, on the other hand, extracting a certain amount of data from them?* How is this possible?

A possible answer to this question must begin by considering how the objects that are the main concern of systematic reviewing acquired such peculiar characteristics. Why are they imbued with enticements and rhetorical forces and what is their role in medical science? A

look at social and historical research on the scientific article reveals that the art of persuasion is integral to the construction of scientific knowledge. Furthermore, as a distinct genre of text, scientific articles were central to the establishment of what came to be known as modern science and the concurrent development of new forms of social regulation. This research suggests that, in modern forms of debate and discussion, the public display of matters of fact, independent from any interested views or opinions, became essential to guarantee an important degree of social cohesion and orderly, gentlemanly discussion. The textual description of these controlled events became, as Steven Shapin has famously put it, 'the expository means by which matters of fact were established and assent mobilised' (Shapin 1984: 484). This textual relation between the laboratory and the audience concerned the *establishment* of fact in the way that it rhetorically enabled the reader to witness the experimental event (Shapin and Schaffer 1985). In this context, research papers can be seen as *rhetorical machines*, deployed in the laboratory's 'organisation of persuasion through literary inscription' (Latour and Woolgar 1986).

The development of these forms of textual persuasion have however been progressively changing within particular disciplines. While 17th or 18th century scientific literature concentrated its rhetorical strategy on demonstrating the author's detachment from its community (of interests, opinion, etc.), papers produced in the 19th century became increasingly embedded in the literature of the field (Brazerman 1988). In this context, the main rhetorical strategy ceases to be creating 'virtual witnesses' of a fact and concentrates on legitimating the claim in relation to shared references of theory and method. In medicine, such strategy is deployed as a means to further enhance the *textual impression* of the authors' detachment from the outcome of the experiment (Kaptchuk 1998). Producing a convincing medical fact entails reporting the experiment by both accounting for the relationships between an array of subjects or entities enrolled in the experiment and the methodological procedures and/or statistical reasoning that regulated the experiment. This convention became particularly important in the reporting of clinical trials, where the demands to control and account for biases (selection, comparator or observer) have been consistently enforced by therapeutic reformers in both sides of the Atlantic (Marks 1997, Mathews 1995).

As papers arrive at the systematic reviewers' office we observe a *confrontation* between these rhetorical forces incorporated in the articles and the data-mining aims of the reviewers. The control that is attempted by the authors of the text over the readers' evaluation of any particular medical technology becomes, in the reviewers' office, a force that is necessary to address in practical and technological terms. Seen from this perspective, the almost Fordist aspect to the reviewers' management of research papers patent in the first fieldnote given above appears in a new light. For it can be argued that at the core of systematic reviewing, is an attempt to neutralise the powers incorporated in texts by their authors. To be able to do this, it is necessary to use devices such as the 'selection criteria', or tabled comparisons between similar studies. In the next section, I identify and describe the devices and practices used by systematic reviewers to extract data from medical papers while making, as another reviewer put it, 'the authors' interpretations of results irrelevant' [Worklog Theresa, 02.03]. To understand this process, the focus will be on the data-abstracting practices, with special attention to how its setting and materials articulate between the parallel procedures of disentanglement and qualification.

The machinery of textual resistance: making texts 'docile'

I start this section with one of my most successful attempts at describing data abstraction:

> DATA ABSTRACTION: situation in which reviewer reads a paper 'through' the protocol [as translated into the data abstracting template], in which s/he looks for items of information in the paper as and if required in the data abstracting template sheet in front of him/her. The reviewer enters data in text template designed with basis on the protocol, using a highlighter pen to direct perception onto certain sentences and numbers and to concentrate attention on sections of the paper. Often reviewers use calculators in order to extrapolate data asked in the protocol but not calculated in the papers themselves.
>
> [Fieldnote #134 Extract – 12/02]

In this fieldnote, I describe the practice of data abstracting by emphasising the way in which reviewers actively seek and construct the information they require. It was clear to me that reviewers were not reading texts as they were displayed on the page, *i.e.* as a linear sequence of sentences. They would often start by reading the 'methods section', to obtain information about the quality of the trial and patient characteristics, and then moved to the result tables. In this, the 'protocol' was central, as an instrument to 'read with'. The importance of this finding was confirmed by the attachment reviewers have to the 'protocol' and the templates they can derive from it. In an interview with Vincent, for example, one of the systematic reviewers, he describes his method of data abstracting thus:

> [My method is] paper sat right underneath my screen, of my PC, ledged [?] on the keyboard and have the pages [of the paper opened] and the abstraction template opened on the PC. [What you do is] you take the trial name, you describe it in er succinct, efficient manner, you discuss, you abstract the quality data, the blinding, randomisation, etc., then you mention, then you start going on more to the numerical data, the baseline data, then the outcomes data . . .
> (. . .)
>
> [Interview Vincent 11/02/03]

In this account, data abstraction is described as a spatial arrangement between reviewer, paper, computer screen and abstracting template. Such an arrangement serves as the setting for an ordered procedure of data retrieval and composition (succinct trial description, etc.). Grounding this layering of the reviewers' knowledge work, is the spatial arrangement of template, screen and paper explained by Vincent above. By arranging his work setting in such a way that the template is its main work object, Vincent could describe his engagement with the text as a structured set of procedures: trial name, its objective, quality of data, etc. In this process, the template plays a fundamental role: it interposes between the text and the reviewers, orientating his/her involvement away from the text as a flow of sentences.

Abstracting as rewriting

Prescribing the order of the procedure described above is the abstracting template. This is an ordinary-looking table drawn in a commonly used word processing package, divided in columns, into which reviewers write the data. Derived from shared standardised protocol forms, these tables usually request the following:

- Trial identification: names of the trial being abstracted, references of the paper(s) from which it was abstracted, the class of the health interventions being evaluated, and the name of the reviewer;
- Comparison: describes the interventions/placebo each arm of the trial is subjected to, identifies the aim of the trial, and arranges the data according to databases;

- Patient characteristics: describes the characteristics of the patients included in the trial;
- Methodological quality: describes how methodological procedures, such as randomisation are accounted for in the paper;
- Comparability: assesses baseline comparability of subject through percentage of age, ethnic groups, and gender;
- Health measurements observed in subjects at the start of the trial;
- End points: gives numerical evaluation of percentage of different health events/measurements in control and intervention arm at the end of involvement in the treatment.

While the identification of the trial seems a quite straightforward piece of information to retrieve from the paper, it is immersed in a complex interaction between part and whole in a process echoing Garfinkel's documentary method of interpretation (Garfinkel 1967, Lynch and Bogen 1996). The recording of reconciliation sessions demonstrates that in accounting for the reporting of methodological procedures reviewers judge the adequacy of the reporting in view of an overall impression of the trial that they construct as they fill in the abstraction template. The template thus allows for the identification of consistencies within and across trials, which effectively supports the formation of this holistic judgement. Another aspect of this structured procedure of filling in the template, is the work reviewers employ in adapting the information the template requires and the reviewer wants, and that being reported commonly in the papers. Such adaptation commonly requires transforming the figures found in the papers to obtain a standardised data form across trials. As Vincent himself explains:

> It's not a simple case of looking at it and seeing what someone's got and typing it down and going away […] we have our own types of data and it's there hidden between the words, camouflaged and everything and we have to use the protocol and other means that we have at our disposal to gather that and at the end of the day that data that we gather it's what is going to be used in the analysis […]
>
> [Interview Vincent 11/02/03]

The adaptation of data requires the use of these 'other means', a variety of instruments of diverse technological design to visually stabilise and easily convert segments of text. In this, highlighter pen markings are used to guarantee that the reviewer's perception of particular portions of the text can easily be sustained. Marginal notes mark the place where particular items of data can be found. Occasional re-calculations of those items can also be found in the papers. These marking on the paper, made as the reviewer fills the template 'spaces', introduce a layer of graphic markings of the reviewers' reading of the paper (Goodwin 1994). This enables an important aspect of data abstraction, its 'detective work' [Interview Sofia 27.02.03]. Having constructed 'a screen' from which to read the paper, the reviewer can safely look for data without danger of becoming 'contaminated by the authors' interpretations of the results'. This method enables the reviewer to go beyond the words, to look for information that is hidden in shady corners of texts: left in footnotes, relegated to the acknowledgements, or in derivate results. The reviewer picks out this information and actively transforms it into data. Through this process, the paper loses its material appearance of a well ordered, sequential set of arguments, turning it into a figuration of interrelated, graphically linked marks and comments. In a more than superficial way, they *rewrite* the paper; and such rewriting is an integral part of the work of re-calculation that is done while abstracting.

The importance of the linking between re-writing and re-calculation of figures in the paper as a means of obtaining data pervaded one of the various training sessions in which

I participated. In one such session we were asked to find data and calculate the 'standard deviation of difference' – SD(D) – for each of the papers we had been given. A fundamental item in this equation was the p value, which was not found *as such* in any of the papers. We looked for the figures that were necessary for a conversion into a p value (standard error, etc.) by scanning the papers looking for clues of where to find such figures and marked them on the paper before proceeding to calculate p value and then the SD(D) for each of the papers. Part of this exercise had been set up by our trainer to illustrate how the data required to complete an equation are often not found as a neatly, discrete figure waiting to be detected. The figures used to construct the value in question had to be gathered from different sections of the paper. In slotting them in the equation, we were effectively assisted by the ballpen circles made around the relevant figures. In this calculating procedure, we looked only for the circles we had made, and immediately discarded the paper to the side. Our rewriting of the paper was fundamental to our focus on and calculation of the p value.

The result of this calculation was then used to calculate the 'standard deviation of difference' for all the other studies in question. In an operation that I recognised as characteristic of statistical reasoning, the extension of one p value to all the other studies was justified on the grounds that we were using the p value chosen for one study as a template of p value setting for studies measuring the same phenomenon. This was furthermore substantiated by the fact that the p value chosen for the study in question seemed to lie within the range of values suitable for testing the null hypothesis. With this value inscribed in all the equations for SD(D), the differences between the results of the studies could be made visible, and compared. The appearance of those differences allowed in turn for the use of another sort of judgement, as it became possible to contrast and evaluate the outcomes of the studies.

The loop in the machinery

In training sessions as in the reviewers' office, the 'protocol' and its templates appear central to the 'epistemic machinery' (Knorr-Cetina 1999) of systematic reviewing, because they encapsulate how data abstracting constructs its empirical referent – the data – through the use of specific instruments of differentiation. Such differentiation is a layered process combining the graphic selection of segments of texts and the constant re-calculation of the figures. In this process, the reviewers' 'rendering practices' – their representation of the text as data – are in close relation with the pre-representational chains that make texts 'docile' (Lynch 1985b). The achievement of this representation of the text, and of the accumulation of similar representation in a database, is linked to the pre-representational work of manipulation, marking, graphic saturation, etc. conducted upon the text.

The tools and arrangements used by the reviewers configure the cognitive involvement of the practice of reviewing, shifting its primary object from the text to the construction and judgment of differences between data of different origins. The possibility of this type of judgement is inextricably linked with the use of platforms, and graphical techniques such as the writing of equations in a white board, or the drawing of 'forest plots' of results of different studies, a picture used iconically as the logo for the Cochrane Collaboration. The results of the studies thus appear as within a set of relations different from their original interpretation and the claims of their authors. They are re-qualified. In this, they might gain new properties, an unimportant study gaining a new significant role in the understanding of an health intervention, or they might lose some of their power. This parallel process of disentanglement and qualification would not be adequately portrayed as a progressive, linear trajectory of action. In the example given in the previous section, for example, it is interesting to note how in calculating SD(D) our orientation was structured by the requirements

of the equations. To re-qualify the papers' claims, we were assisted by two equations, which acted as boundary objects (Leigh-Star and Griesemer 1989) articulating the disentanglement/qualification process in a non-linear way.

The complex ways in which these articulations work can be seen in how they mediate across the diverse sets of activities within systematic reviewing and might inclusively entail the protocol itself changing its requirements and conditions. This double complexity can be illustrated by a fieldnote:

> When I entered the systematic reviewers' office, I immediately understood, through Vincent's expression, that there was something wrong. It had become apparent, by comparing their results with similar studies, that their meta-analysis had missed 'a very important trial'. This had been because the selection criteria they had set up were too stringent, and the trial in question did not have the required number of subjects in the study. The trial was however considered to be the standard of research within the specific medical speciality and all previous reviews of the same class of drugs had taken it into consideration. Vincent reported that George, the unit's director, said that excluding the trial would most likely be taken as a confrontation by lead researchers in the area. Including the trial would entail running a new literature search, and Vincent guessed that that would mean abstracting another 25 papers and re-calculating 'the whole thing'. Most importantly, it would most probably result in making their meta-analysis much similar to previous ones. He reasoned that if a new search and review was to be conducted with new inclusion criteria for number of subjects, some other criteria in the protocol would have to be made more rigorous.
>
> [Fieldnote #52, 11/02]

The new properties that statistical results gain as an outcome of the process of disentanglement and qualification in systematic reviewing depends on this ability of the meta-analysis to be itself compared to similar studies. This means that it is important to be able to adjust the various elements of the review – inclusion criteria, abstracting templates, statistical models, etc. – during the course of research. The maleability of the protocol is thus fundamental to the practice of systematic reviewing because it allows for 'new ideas and new slants (to) keep coming in' [Maria 2.03.03, Interview] and increases the reviewers' ability to produce new knowledge claims.

In this instance, the loss of value that came from having to use the same inclusion criteria as previous studies was compensated by the choice of the statistical model that underpinned their claims. While most of the previous meta-analyses of the health intervention they were studying used a 'fixed effects model', Laura, the team's head statistician, decided to use a 'random effects' model. These refer to the theory and method of calculation of the aggregated effects of different studies: 'a fixed effects model assumes that there is one – and only one – real treatment effect, which applies to all populations', while 'a random effects model assumes that the treatment effects in different populations vary a bit (although they will be correlated) because the populations differ. Technically, we assume that the treatment effects in different populations come from a normal distribution.' [Laura, personal communication]. By applying this model and by comparison with the meta-analyses that had become potential competitors, the team was able to produce a different, re-qualified calculation of the 'most probable real effect' of the health intervention. Also, as it turned out, this calculation technique allowed the team to show that the estimated treatment effect reported in the trial that had made them change their inclusion criteria was on the edge of the acceptable range within the normal distribution curve. The original inclusion criteria

were subsequently recovered as part of an ideal protocol to inform further design of randomised controlled clinical trials in this area.

In this, systematic reviewers' aims resemble those of the authors of the texts that they have been attempting to deconstruct and re-qualify, as they too organise their rhetorical devices in the form of presentations, reports and papers. The non-linear, looping aspect of the process of articulation of disentanglement and qualification seems thus to be linked with the credibility that reviewers seek to obtain from their peers, including clinical trial designers. Such peer-based credibility is, however, intimately linked with the strength systematic reviewers can bestow on their analyses when addressing other constituencies. In fact, for meta-analyses and systematic reviews to circulate outside the laboratory, peer-based credibility has to be carefully combined with political awareness.

Out of the textual laboratory

In constructing their analyses, systematic reviewers are required to *entangle* their calculations with the technical devices and the social and political relations that the devices mediate. The importance of this politico-technical entanglement is demonstrated in the ways in which meta-analyses are circulated outside the 'laboratory' in, for example, clinical practice guideline meetings. It can, in fact, be argued that these meetings constitute a fundamental test of the persuasive strength of clinical evidence (Moreira 2005), and can thus be used as model for the political circulation of systematic reviews.

In these meetings meta-analytical results are tabulated, summarised and discussed by multidisciplinary inter-professional groups aiming to produce recommendations for clinical practice for a particular constituency or constituencies. In their political evaluation, reviewers, methodologists and healthcare practitioners, presented with the results of a meta-analysis, assess how the results of the review might interfere with the distributions of power and accountability within healthcare institutions. In the extract below, group members are trying to decide on how to account for the exclusion of a trial from their review that has become publicly prominent meanwhile. Although portraying a predicament about exclusion of trials that might seem similar to the one described in the previous section, in this instance group members are more concerned with assessing how the persuasive strength of the reasons for excluding the trial will measure against a state of affairs where that trial has gained public recognition. They are thus concerned with evaluating the *political robustness* of that decision:

> Consultant 2: I think that if we had decided to include the [Study X] within the [meta-analysis] that we discussed we should have [reached different conclusions], as even [methodologist] has said. I think [consultant 1] and I have said why we think that the trial does not show specific benefits for [drug y]. In one sense it's easier to kind of put it to one side. The problem though is that that trial has just had one of the most successful marketing campaigns ever and most people in primary and secondary care [believe that there are benefits beyond those of other drugs], and it's got in the [national tabloid] because [they] have been extremely clever with their marketing and PR. So the question is whether a guideline should attempt to address [this] major misconception in most doctors in the UK [but] I don't have an answer to that. I think it's difficult for a guideline to start to get engaged with

that because it's kind of almost … […] [But] personally I am in favour of some comments on [Study X] actually in the major text.

Methodologist: But there's a comment on why it's not included.

Consultant 2: But I would also make a comment as to why we [did not include it].

Methodologist: Well, I could put another sentence in here and say there's been recent public controversy about whether [this statistical result] is correct for [the patient] group [we are concerned with].

[…]

Methodologist: [But] let's just wait [and discuss this again]: in practice, these treatments were at least third down the line on our recommendations.

GP 1: I think that where the drug company's placing it is quite different from where the evidence places it. But if you feel that that's absolutely clear . . .

Methodologist: I think actually it's something for when the guideline is being [reviewed by the stakeholders]. If we adapt it [in line with the guidelines of medical society], it's something that will be noticed and […] I think the risks of straying into it and us giving it pre-eminence are greater than [leaving it as is] . . .

[Group1/Mtg10/48-49/edited transcript]

In the extract, consultant 2 starts out by aligning his and consultant 1's positions with the methodologist's decision to exclude Study X from the meta-analysis presented to the group, only then to lay out the public opinion landscape against which such a decision will be read: a powerful marketing campaign, news in a national tabloid, 'most people' believing in the importance of results of Study X. According to him, the group is faced with a peculiar dilemma: not engaging with a difficult, convoluted political issue, or – his personal favourite – facing it upfront 'in the major text'. The methodologist's response is that this upfront approach would give too much importance to a trial that is only relevant for treatments 'down the line in [their] recommendations'. GP 1 reminds the methodologist that this position is likely to be met with disagreement by the drug company that sponsored the trial. The methodologist insists on the value of his low-profile approach, preferring to model the group's position with an established medical society.

The group's disagreement is about whether or not to confront the configuration of actors that constructed the public recognition of the treatment being discussed. The chances of the guideline being publicly accepted depend partially on whether their statements on evidence are aligned or dis-aligned with a variety of actors. By excluding a trial – the conclusions of which have been circulating in the media and health-policy debates – from their meta-analysis, they were effectively disengaging from a particular configuration of actors. Consultant 2 seems to think that, on the strengths of the meta-analysis, the group alone can stand against this set of actors by itself. The methodologist, on the other hand, prefers to downplay the divergence with the drug company, and highlight which actors can be drawn into the group's positioning if necessary.

Through this example, it is possible to understand the predicaments with which meta-analyses are faced *in order to* leave the laboratory and to circulate in health-policy debates. In this, it is key to assess how its statistical calculations might interfere with the distribution of actors within the healthcare arena. In so doing, actors are usually faced with having to choose between two or more sets of actors. This choice implies evaluating how best to *disentangle* from one set of actors as well as devising strategies to become associated with others. On the one hand, it is necessary to evaluate collaboratively how to express

disagreement, while on the other, match this approach with the group's strategy of alignment. The political *qualification* of a systematic review implies weaving together its calculations and the political distributions that best fit them.

Their political trajectory however also depends on how these political configurations are supported by the fit they attain with meta-analytic calculations. In this, as was argued in the beginning of the paper, systematic reviews are valued for their role in maintaining an appropriate boundary between biomedical science and health politics. This trend has been reinforced, in Britain at least, by recent government initiatives that underline the importance of 'evidence' in policy making. It is known that within the policy arena 'evidence' is reinterpreted through a pragmatic political lens that manages expectations and public opinion as well as resources (Harrison 1996, May 2005). In this, the political constituencies mediated by systematic reviewing seem to expect that such a boundary can be stabilised by emphasising its disentangling ability and suppressing its qualification/entangling procedures. By accentuating how systematic reviews and meta-analyses construct knowledge by disentangling their data from misleading papers written by 'interested', privately-funded and/or career-driven authors (Chalmers 2001), political actors invest in deploying the 'public face of systematic reviews'. The research presented here suggests, however, that such disentaglement depends as much upon a practice of qualification that ties systematic reviews to mundane devices, graphical techniques and political constituencies alike.

Conclusion

Two decades ago, Phil Strong predicted that the academic disciplines that successfully encroached in clinical medicine would 'depend as much on the wider success of their patrons as on their own academic effort' (Strong 1984: 358). By the turn of the century, medical statistics and the methodologies of health-technology assessment have effectively been able to gather the social and political sponsorship to influence health policy and some sectors of clinical medicine. This process has positioned systematic reviewing in the pivotal role of mediating between science and politics in healthcare. The research presented in this paper, however, suggests that the academic efforts developed within systematic reviewing to address such demands might nevertheless be in tension with the expectations that the social and political constituencies put in these research practices.

The paper has argued that knowledge construction in secondary research in healthcare is structured upon parallel attempts a) to extract data from the milieus in which they are commonly found (databases, texts, research centres, etc.) and b) to re-qualify the value of those data through a series of comparisons across a variety of 'platforms' (tables, graphs, equations, political controversies, etc.). Fieldwork data suggest that these simultaneous strategies of disentanglement and qualification structure the sets of activities into which systematic reviews and meta-analyses are ordinarily divided – data selection, data abstraction and calculation – in a continuous, dynamic interrelation between mutually dependent locally situated activities. This relationship between these two processes relies on the non-linearity of the practice of reviewing and re-calculating. Such a description of systematic reviews is furthermore consistent with studies of other forms of calculation such as economics (Callon 1998b, Callon and Muniesa 2002). As in economics, the accomplishment of these two processes is fundamental in establishing the persuasive power of systematic reviews and *equipping political action* in contemporary healthcare.

New research is, however, necessary to understand how systematic reviews are reconfigured within the practices of health politics itself. Such research endeavour should be an empirical,

preferably ethnographic, study of the politics as a technical activity (Barry 2001). This will be important material to shape a discussion centred on the types and intensity of the entanglements that might legitimately mediate between biomedical science and health politics.

Address for correspondence: Tiago Moreira, School of Applied Social Sciences, Durham University, 32 Old Elvet, Durham, DH1 3HN
e-mail: tiago.moreira@durham.ac.uk

Acknowledgements

I am indebted to the staff at the unit where this research was conducted for their patience, criticisms and encouragement. The paper also profited from the kind suggestions of Pascale Bourret, Michel Callon, Neil Jenkins, Carl May, Maggie Mort and Tim Rapley and two anonymous referees. It benefited from comments of participants at the *Organization Analysis Informing Social and Global Development* – 19th EGOS Colloquium, Copenhagen, 3rd–5th July 2003, at a 'lunch time seminar' at the Health Research Institute, Lancaster University, 11th December 2003, and the workshop 'Saúde, Natureza e Peritagem', Instituto Superior de Economia e Gestão, Lisbon, 7–8 June 2004.

References

Abel-Smith, B. (1996) The escalation of health care costs: how did we get there? In OECD (ed.) *Health Care Reform – the Will to Change*. 1st Edition, 8, 17–30. Paris: OECD.

Ashmore, M., Mulkay, M. and Pinch, T. (1990) *Health and Efficiency: a Sociology of Health Economics*. Milton Keynes: Open University Press.

Barry, A. (2001) *Political Machines: Governing a Technological Society*. London: Continuum.

Brazerman, C. (1988) *Shaping Written Knowledge*. Madison: University of Wisconsin Press.

Callon, M., Lascoumes, P. and Barthe, Y. (2001) *Agir dans un monde uncertain: essai sur la democratie technique*. Paris: Seuil.

Callon, M. (1998a) Introduction: the embeddedness of economic markets in economics. In Callon, M. (ed.) *The Laws of the Markets (Sociological Review Monograph)*. Oxford: Blackwell.

Callon, M. (1998b) An essay on framing and overflowing: economic externalities revisited by sociology. In Callon, M. (ed.) *The Laws of the Markets*. Oxford: Blackwell.

Callon, M. and Muniesa, F. (2002) Economic markets as calculative and calculated collective devices. New York Conference on Social Studies of Finance, Columbia University. Ref Type: Unpublished Work.

Chalmers, I. (2001) Comparing like with like: some historical milestones in the evolution of methods to create unbiased comparison groups in therapeutic experiemts, *International Journal of Epidemiology*, 30, 1156–64.

Chalmers, I., Hedges, L.V. and Cooper, H. (2002) A brief history of research synthesis, *Evaluation and the Health Professions*, 25, 1, 12–37.

Egger, M., Davey Smith, G. and Altman, D.G. (2001) *Systematic Reviews in Health Care: Meta-analysis in Context*. London: BMJ Publishing Group.

Egger, M., Davey Smith, G. and O'Rourke, T. (2001) Rationale, potentials and promise of systematic reviews. In Egger, M., Davey Smith, G. and Altman, D.G. (eds) *Systematic Reviews in Health Care*. London: BMJ.

Garfinkel, H. (1967) *Studies in Ethnomethodology*, 2nd Edition (1984). Cambridge: Polity Press.

Ghali, W.A. and Sargious, P.M. (2002) The evolving paradigm of evidence-based medicine, *Journal of Evaluation in Clinical Practice*, 8, 2, 109–12.

Gieryn, T.F. (1999) *The Cultural Boundaries of Science*. Chicago: University of Chicago Press.

Goodwin, C. (1994) Professional vision, *American Anthropologist*, 96, 3, 606–33.
Guston, D.H. (2001) Boundary organizations in environmental policy and science: an introduction, *Science Technology and Human Values*, 26, 4, 399–408.
Harrison, S. (1998) The politics of evidence-based medicine in the United Kingdom, *Policy and Politics*, 26, 1, 15–31.
Harrison, S. (2001) The corporate practice of medicine: competition and innovation in health care, *Sociology of Health and Illness*, 23, 1, 133–4.
Harrison, S. (1996) The politics of evidence-based medicine in the United Kingdom, *Policy and Politics*, 26, 15–31.
Harrison, S. (1999) Clinical autonomy and health policy: past and futures. In Exworthy, M. and Halford, S. (eds) *Professionals and the New Managerialism in the Public Sector*. Buckingham: Open University Press.
Hunt, M. (1997) *How Science Takes Stock: The Story of Meta-Analysis*. New York: Russell Sage Foundation.
Irwin, A. and Michael, M. (2003) *Science, Social Theory and Public Knowledge*. Milton Keynes: Open University Press.
Kaptchuk, T.J. (1998) Intentional ignorance: a history of blind assessment and placebo controls in medicine, *Bulletin of the History of Medicine*, 72, 3, 389–433.
Katz, J. (2002) From how to why: on luminous description and causal inference in ethnography (Part I), *Ethnography*, 2, 4, 465.
Knorr-Cetina, K. (1981) *The Manufacture of Knowledge: an Essay on the Constructivist and Contextual Nature of Science*. Oxford: Pergamon Press.
Knorr-Cetina, K. (1999) *Epistemic Cultures. How the Sciences Make Knowledge.* Cambridge MA: Harvard University Press.
Latour, B. (1990) Drawing things together. In Lynch, M. and Woolgar, S. (eds) *Representation in Scientific Practice.* Cambridge MA: MIT Press.
Latour, B. and Woolgar, S. (1986/1979) *Laboratory Life. The Construction of Scientific Facts*, 2nd Edition. Princeton: Princeton University Press.
Lehoux, P. and Blume, S. (2000) Technology assessment and the sociopolitics of health technologies, *Journal of Health Politics Policy and Law*, 25, 6, 1083–120.
Leigh-Star, S. and Griesemer, J. (1989) Institutional ecology, 'translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39, *Social Studies of Science*, 19, 387–420.
Levin, A. (2001) The Cochrane Collaboration, *Annals of Internal Medicine*, 135, 4, 309–12.
Light, D. (2000) The origins and rise of managed care. In Phil, B. (ed.) *Perspectives in Medical Sociology*. Prospect Heights Ill: Waveland Press.
Light, D. and Levine, S. (1988) The changing character of the medical profession: a theoretical overview, *Milbank Quarterly*, 66, 10–32.
Lynch, M. (1985a) *Art and Artifact in the Laboratory: a Study of Shop Work and Shop Talk in a Research Laboratory*. London: Routledge and Kegan Paul.
Lynch, M. (1985b) Discipline and the material form of images: an analysis of scientific visibility, *Social Studies of Science*, 15, 1, 37–66.
Lynch, M. and Bogen, D. (1996) *The Spectacle of History. Speech, Text, and Memory at the Iran-Contra Hearings*. Durham: Duke University Press.
Marks, H.M. (1997) *The Progress of Experiment*. New York: Cambridge University Press.
Marks, H.M. (1999) The emergence of statistical proof in medicine, *Recherche*, 316, 76–81.
Mathews, J.R. (1995) *Quantification and the Quest for Medical Certainty*. Princeton NJ: Princeton University Press.
May, C. (2006) Mobilising modern facts: health technology assessment and the politics of evidence, *Sociology of Health and Illness*, 28, 5, 513–32.
Mckeown, T. (1979) *The Role of Medicine: Dream, Mirage or Nemesis?* Oxford: Blackwell.
Moreira, T. (2005) Diversity in clinical guidelines: the role of repertoires of evaluation, *Social Science and Medicine*, 60, 1975–85.

Mulrow, C.D. (1994) Rational for systematic reviews, *British Medical Journal*, 309, 597–8.
Porter, T.M. (1995) *Trust in Numbers: the Pursuit of Objectivity in Science and Public Life*. Princeton NJ: Princeton University Press.
Rawlins, M. (1999) In pursuit of quality: The National Institute for Clinical Excellence, *Lancet*, 353, 9158, 1079–82.
Roberts, I. and Schierhout, G. (1997) The private life of systematic reviews, *British Medical Journal*, 315, 7109, 686–7.
Shapin, S. (1984) Pump and circumstance: Robert Boyle's literary technology, *Social Studies of Science*, 14, 481–520.
Shapin, S. and Schaffer, S. (1985) *Leviathan and the Air-Pump. Hobbes, Boyle and Experimental Life*. Princeton NT: Princeton University Press.
Strauss, A. and Corbin, J. (1998) *Basics of Qualitative Research. Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks CA: Sage.
Strong, P.M. (1984) Viewpoint: the academic encirclement of medicine? *Sociology of Health and Illness*, 6, 3, 339–58.
Tanenbaum, S.J. (1994) Knowing and acting in medical research: the epistemological politics of outcomes research, *Journal of Health Politics, Policy and Law*, 19, 27–44.
Timmermans, S. and Berg, M. (2003) *The Gold Standard: the challenge of EBM and Standardization in Medicine*. Philadelphia: Temple Univeristy Press.
Yusuf, S., Peto, L., Collins, R. and Sleight, P. (1985) Beta blockade during and after myocardial infarction. An overview of the randomized trials, *Progress in Cardiovascular Diseases*, XXVII, 5, 335–71.