Research and Theory for Nursing and Midwifery: Rethinking the Nature of Evidence

Christine McCourt, BA, PhD, Reader, Thames Valley University, UK

Address correspondence to:

Christine.McCourt.1@city.ac.uk

Submitted for publication: 10 January 2005
Accepted for publication: 24 February 2005

Keywords: Evidence, Complexity, Systems, Evidence-Based Medicine, Methodology, Epistemology, Midwifery

ABSTRACT

Background and rationale

The rise in the principles of evidence-based medicine in the 1990s heralded a re-emerging orthodoxy in research methodologies. The view of the randomised controlled trial (RCT) as a “gold standard” for evaluation of medical interventions has extended recently to evaluation of organisational forms and reforms and of change in complex systems—within health care and in other human services. Relatively little attention has been given to the epistemological assumptions underlying such a hierarchy of research evidence.

Aims and methods

Case studies from research in maternity care are used in this article to describe problems and limitations encountered in using RCTs to evaluate some recent policy-driven and consumer-oriented developments. These are discussed in relation to theory of knowledge and the epistemological assumptions, or paradigms, underpinning health services research. The aim in this discussion is not to advocate, or to reject, particular approaches to research but to advocate a more open and critical engagement with questions about the nature of evidence.

Findings and discussion

Experimental approaches are of considerable value in investigating deterministic and probabilistic cause and effect relationships, and in testing often well-established but unevaluated technologies. However, little attention has been paid to contextual and cultural factors in the effects of interventions, in the culturally constructed nature of research questions themselves, or of the data on which much research is based. More complex, and less linear, approaches to methodology are needed to address these issues. A simple hierarchical approach does not represent the complexity of evidence well and should move toward a more cyclical view of knowledge development.

INTRODUCTION

This article does not set out to provide a critique of methodologies currently favoured in evidence-based health care or policy, particularly the randomised controlled design. It does, however, advocate a critical evaluation of such approaches to research, their limitations, and the ways in which they can be most productively applied to the more complex interventions that are often the subject of research in health care. The critical perspective advocated is not simply a technical matter of highlighting and resolving the challenges of designing good experiments in complex situations but is also informed by theoretical and philosophical considerations about epistemology and the nature of scientific evidence.

The view of the randomised controlled trial as the gold standard in health research reflects its robustness as a design that can minimise certain systematic and unanticipated biases in research samples, to enable a fair comparison or test of a new or established intervention. It is unfortunate that, in practice, the use of such terminology as gold standard may be interpreted as implying that other research designs are second rate, or lacking in rigour, rather than complementary and appropriate for different research contexts, stages, and interests (Glasziou et al. 2004).

As Oakley (1992, 1998) has eloquently reminded us, evaluation of social, educational, and health policy and practice and the use of experimental methods in such evaluations have a long and distinguished history showing

that many policies that intuitively promised benefits, when rigorously evaluated, did not show significant benefits and sometimes had negative effects. Nonetheless, there are particular challenges in applying experimental methods to complex social or organisational contexts, which need to be recognised in planning and evaluating research (Medical Research Council [MRC] 2000).

Kuhn's (1970) study of scientific paradigms argued that much of scientific enquiry has not followed the logic of inductive theory development followed by rigorous theory testing that Popper's (1959) principle of falsifiability suggests. People perceive and interpret the data involved in research selectively, and this constitution of the field of evidence is something that is not necessarily explicit or self-conscious. It is culturally framed and mediated (Clifford & Marcus 1986). Thus, the very questions researchers seek to address, the ways in which questions are framed, notions of relevance, what is viewed as data, and a primary or secondary outcome measure are influenced by prevailing epistemologies. Principles of measurement themselves are similarly historically and culturally framed and situated (Foucault 1973; Arney 1982; Oakley 1992; Thomas 1992). This suggests that although scientific method (and that is something I treat as equally applicable to all research, including the use of qualitative methods) should clearly reflect the principle of falsifiability, it should equally reflect the ways research questions are constituted and the understanding of what data or evidence is as well as from whose viewpoint.

This article illustrates the need for caution and some of the possible limitations in using experimental designs by reference to several RCTs of midwifery and nursing practices. Researching nursing and midwifery practices can present particular challenges since they often involve complex settings and organisational frameworks and the nature of the intervention itself may not be readily amenable to definition, let alone control. It is argued that the current prevailing structures for commissioning, reviewing, and disseminating “evidence-based” policy or services have not yet satisfactorily acknowledged and integrated these challenges, which are more than a matter of technical improvements in trial design and quality.

BACKGROUND

Evidence-based medicine as a movement?

The nurse and sociologist Traynor (2002, 2003), drawing on the work of anthropologists such as Douglas (1966), has drawn an analogy between the late 20th century drive toward evidence-based medicine (EBM) and a new religious movement: It showed evangelistic features and had charismatic leadership, with clear insiders and outsiders (the unconverted). To take this analogy one step further, analysis of the structures and processes of new religious movements has demonstrated that an early period of charismatic leadership and small, committed following (believers) is followed by a period of routinisation and orthodoxy as the movement becomes more widely accepted and institutionally developed (Worsley 1957; Barker 1989). Similar arguments have been applied to EBM as a social movement (Pope 2003).

Such analyses attempt to understand EBM as a part of a historical and cultural movement and so to support a questioning view.

It is at a later stage in the development of social movements that a more critical evaluation is likely to be developed and tolerated. This is evidenced here in the recent publication of critical articles in the medical press on the concept of evidence for health care (Goodman 1999; Greenhalgh 1999; Black 2001).

Theoretical roots

Taking the theme of EBM as a social movement suggests that it is important to understand its socio-cultural as well as theoretical roots—and consider the assumptions on which it has operated. In doing so, my own theoretical perspective includes the argument that cognition is related to both environment and ideology: While there are underlying biological or physiological systems that structure human cognition universally, it is also culturally shaped to a considerable degree. Cultural systems of belief, or ideology, are written on the individual, social, and political body and interact iteratively with structure and environment to shape systems of belief and action (Bourdieu 1972; Scheper-Hughes & Lock 1987; Bloch 1989).

Experimental research has long historical roots but is particularly associated with the enlightenment and the rise of positivist science in Europe, which sought universal laws and prediction to explain the natural and social world. In an era of rapid socio-economic and cultural change, a combination of observational and experimental research underpinned much of scientific

development of the modern era. The shift from Newtonian to quantum theory in physics in the 20th century challenged positivist science in suggesting that the validity of physical theories may be influenced by the scale and context of operation and the viewpoint, or even presence, of the observer (Capra 1983). Thus developments in the natural as well as the social sciences encouraged the development of post-modern and critical theory.

Experimentation as represented in the randomised controlled trial in health care uses a positivist theoretical framework and likewise has truth value (or is effective—like Newtonian physics) within certain parameters. Its value depends on a level of control that is not always achievable, despite the best efforts of researchers, and that tends to break down particularly once more complex contexts and interventions are studied. This has been extensively discussed by thinkers in evidence-based medicine, in responding to practitioners' concerns about application of trial evidence to individual cases and to everyday practice (Sackett et al. 1996).

A further problem that may underlie debates about the status of evidence and the "problem" of its implementation is that the research paradigm, of necessity (to achieve control), separates the intervention from its context—professional, socio-cultural, and operational. Researchers and practitioners are thus presented with a problem of how to (re)connect *evidence* with *practice.* An examination of commissioned work on implementation of research evidence suggests that the original theoretical model for this work was also essentially positivist—a rational-legal or bureaucratic model of

action somewhat akin to Weber's ideal types of knowledge and organisation. For example, trials were commissioned that compared different approaches to getting evidence across to practitioners and latterly to patients (Bero et al. 1998). These have shown limited effectiveness, and research commissioners have been forced to reconsider how *evidence* can be made to work in practice when *behaviour* fails to fit the logical-positivist model. Through the case studies in this article, I suggest that a fundamental problem within the current construction of evidence within EBM is that context, decision-making processes, and behaviour have tended to be written out of the model as potential confounders, contaminants, or sources of bias. Such factors are seen as standing in the way of evidence—in research design and in implementation—rather than as part of the system to be studied (Wood et al. 1998; Hawe et al. 2004).

This has been recognised to some degree in the Medical Research Council's (2000) recent guidelines on trials of complex interventions, where these problems are most relevant. It might be argued, however, from the cases reviewed here, that such guidelines should be applied more widely. In the model advocated by the MRC, knowledge development is seen as a cyclical rather than a linear process. Similar arguments have been derived from systems theory and complexity theory (Bateson 1985; Plsek & Greenhalgh 2001; Downe & McCourt 2004). The implications of this for the hierarchy of evidence are that the RCT may be seen as an essential part of a wider approach to scientific knowledge, particularly valuable for certain questions

and stages of the research process, rather than as a gold standard against which all other forms of evidence should be directly compared.

EXAMPLES OF RESEARCH IN MIDWIFERY: A CRITICAL DISCUSSION

Case study 1: An RCT of routine episiotomy for perineal trauma

An early and powerful example of experimental research in maternity care was led by a midwife in the United Kingdom (UK), collaborating with an obstetrician and a social scientist (Sleep et al. 1984). A randomised controlled trial was used to examine the clinical impact of routine episiotomy at a time when most obstetric developments were in increasingly routine use without being evaluated systematically. The study was prompted by the doubts of both women and midwives about the value of routine episiotomy use rather than selective use, and similar trials were then conducted in a wide range of countries (Graham & Davies in press).

This intervention lent itself well to an experimental design since, although it could not be “blinded” (as with drug treatments), there was no disappointment factor for women in not being selected to be in the intervention group and there was sufficient uncertainty about the practice for professionals to respond comfortably to a trial (Edwards et al. 1998). Additionally, it was a routine practice that most midwives had considerable experience and skill in using, rather than an innovative or rare intervention that might be difficult to test pragmatically. The nature of the intervention itself was relatively simple and definable. The outcome measures included women’s short- and long-term self-reports of healing and pain as well as professional assessments of

healing. The trial was thus also able to combine both clinical and qualitative experiential data effectively. It indicated that routine episiotomy had no discernable clinical or other benefits and its use did decrease in the UK in the 1980s (Department of Health 2001-02).

It was interesting to note that the results accorded with women's views and experiences and midwives' concerns, which perhaps accounted for the apparent rapid response to the results. Midwives in the UK were arguably willing to put practices and interventions to the test in this way because of their particular history of scrutiny of practice and due to an ideological and power struggle with obstetrics: They had a professional interest in challenging routine use of episiotomy and this was an intervention within their sphere of influence, rather than highly dependent on inter-professional agreement. This was also a low-technology intervention, without considerable investment in equipment or drugs that might mean loss of a market for stakeholders in the business of care (Pieters 1998). Logically, it might be expected that expensive interventions not shown to be effective would be dropped in a limited health care economy. From the economist's viewpoint, unnecessary interventions would represent an opportunity cost. Experience of the use of a range of technologies in maternity care—for example, the routine use of electronic fetal monitoring or ultrasound—suggests otherwise (Audit Commission 1997).

This case study shows that RCTs can be a powerful and effective way of evaluating health interventions, especially to question and test accepted

wisdoms. However, whether their findings have an impact on practice may depend on far broader factors than those taken into account in a trial.

Case study 2: An RCT to evaluate "soft" interventions and outcomes

My second example is of the use of experimental design to test the effectiveness of a "soft" intervention in midwifery: social support. Oakley and colleagues (1990) undertook a randomised controlled trial of the impact of social support in pregnancy on maternal and infant health. The trial built on a considerable body of observational, epidemiological, psychological, and qualitative research from a range of countries, which had suggested associations of social support with positive health outcomes (Mander 2001; McCourt 2003). This included Oakley's own in-depth qualitative research on women's experiences of childbirth that led her, through more inductive methods, to the hunch that social support was valued by women and could make a difference to outcomes for those who were "at-risk" in some way (Oakley 1980, 1986). Birthweight was chosen as the primary outcome measure, as a relatively reliable indicator of maternal and infant health that was amenable to measurement (Oakley 1992).

Experimental research in this area was a challenge since social support was not so easily definable or controllable as many health interventions, although the trial was centrally managed and the support was provided by a small group of midwives who received common training, communicated closely, and were happy to follow agreed guidelines. Additionally, wide ranging and in-depth qualitative data were gathered to assist in understanding the nature

of the intervention as practised and women's responses to it. A major challenge in mounting a trial of an intervention of this type—particularly one where the terminology itself conveyed common perceptions of being beneficial—was the ethical and methodological problems of recruiting women to a trial who might feel disappointed by random allocation into the control group, given that "perceived" support is thought to be an important aspect of effectiveness of social support (Ross et al. 1999; Mander 2001). Although preference trials have been developed to address such problems, these have also been subject to problems and would unlikely be applicable in the case of all interventions (Edwards et al. 1998). The breadth of data gathered enabled the researchers to reflect on this and did suggest that simply by being part of the trial and being invited to complete questionnaires, many women perceived themselves to have received additional social support. Such findings are often described as the "Hawthorne" effect of research and are particularly pertinent to trials where the intervention is not amenable to blinding in the manner of a drug trial and where, for ethical reasons, participants must be fully aware of the range of interventions they may be allocated to (Edwards et al. 1998).

This study, led by a social scientist committed to the principle of using experimental methods to provide sounder testing of social interventions, did not entirely follow the usual criterion of confining itself to a single predefined outcome. While no statistically significant difference was found in birthweight—the primary outcome measure—significant differences were found in other, less direct health indicators, such as rate of casualty

department visits (Oakley et al. 1996). The researchers hypothesised that the social support had indirect positive health effects in encouraging women to seek out and use other, less formal sources of support more effectively, including their own partners. Clearly, if the researchers had taken the more narrow approach that is preferred statistically, and for systematic review and meta-analysis, for its simplicity and decreased danger of producing spurious positive effects, their conclusions would have been less informative.

Simplicity in research design rests on the capacity to know in advance what the most desirable or most likely risks or benefits will be, as well as having an adequate and reliable system of measurement. One possible reason for the lack of effect on birthweight in this trial was that the team simply did not have adequate knowledge or basis for deciding who should be considered both “at risk” and in need of additional social support. The majority of trials from a wide range of countries looking at support and birthweight have also failed to find significant increases, and those using narrower measures have simply concluded that provision of social support makes no difference to maternal and infant health (Langer et al. 1996). A systematic review for the Cochrane Library (Hodnett & Fredericks 2003) drew similar conclusions, but the reviewers noted the wide range of interventions—all described as social support—involved a variety of contexts and applied to diverse women. Additionally, not all attempted to determine whether the intervention was perceived as supportive by the women themselves. The only trial to have shown significant increases in birthweight, conducted in the United States, was very tightly focused on women who lacked the usual informal sources of

social support and offered peer rather than professional support that modelled the kind of support a mother or partner would give (Norbeck et al. 1996). A further possible limitation of such studies is that the social support intervention offered is limited and of short duration—and so unlikely to make any measurable impact on the very broad systemic effects of social inequality and social conditions or social capital (McCourt 2003). This draws us back to the problem of incorporating social context in conducting and interpreting experimental research.

This case clearly illustrates the difficulty of maintaining experimental levels of control in a complex situation and where participants are aware of their experimental situation. It also highlights the difficulties of using a trial design where the nature and effectiveness of the intervention itself is strongly influenced by personal perceptions and responses, where the social context and manner in which the intervention is provided and perceived is critical, and with the complicating factor of the potentially supportive effect of feeling involved in a study of this type.

Case study 3: Evaluation of MIDIRS (Midwives Information and Resource Service) informed choice leaflets

An RCT was conducted in the UK to evaluate the effects of a set of leaflets designed to support informed choice on women's perceived choices, knowledge, and satisfaction with information, relating to 10 interventions for which good evidence was available. The trial found no difference in knowledge or outcomes between those units randomised to use the leaflets

and those that did not (O'Cathain et al. 2002). An ethnographic study conducted by the authors showed that, despite the trial protocols, the leaflets were not being used as intended for a variety of reasons including staff lack of time and staff attitudes. The trial itself was able to show that the use of "decision aids may not be effective in the real world" (O'Cathain et al. 2002, p643.) but would not have offered any understanding of *why* they may not work in practice, or how to enhance or ensure their effectiveness. In contrast, the ethnographic study showed that a combination of time pressures, midwives rarely discussing the contents of the leaflets, professional perceptions around "right" and "wrong" choices and the nature of litigation, hierarchical power structures where obstetricians defined the norms of practice, and actual lack of availability of choices contributed to their ineffectiveness (Stapleton et al. 2002). The study authors concluded that the conditions of use promoted "informed compliance" rather than informed choice.

This case shows how combining other forms of research—such as ethnographic research—with trials may improve our understanding, particularly of how or why an intervention does or does not have the expected effect. It suggests that far greater attention to other forms of research prior to and during a trial may help to avoid researchers wasting time and money on testing an intervention that does not have a chance to "work" due to contextual and cultural factors. While a pragmatic approach might be to say "if it didn't work in practice, it doesn't work," such an approach to research will not enhance knowledge and may even undermine it, offering misleading

conclusions. If the trial in this case had been conducted without an ethnographic study, the conclusions would simply have been that the informed choice leaflets were ineffective. While, in one sense this was so, the reasons were complex and could not be deduced from an experimental design.

Case study 4: An RCT of nursing support in North America

The epidemiological nurse researcher Hodnett has played a major role in developing and promoting principles of evidence-based health care by conducting influential RCTs and overviews of research on childbirth. Among these were trials of the effects of continuous support in labour and Cochrane overviews of trials that have been referenced here. Klaus and colleagues (1986) and Sosa and colleagues (1980) conducted a series of trials in Guatemala in the 1970s that were able to establish the effectiveness of a support person who stayed with the woman during labour. The Guatemala trials reported highly significant reductions in labour complications, interventions, and duration in women supported by a “doula” (a non-professional but experienced woman providing hands-on physical and emotional support, never leaving the woman alone). Sub-group analysis also indicated these effects were stronger in more socially isolated women. Like Oakley’s social support trial (see case study 2), these studies built on earlier work on effects of social support using a range of disciplines and methodologies. These trials were important to test the hypotheses developed in such earlier work. Later trials developed in North American hospitals showed far more modest effects, and Hodnett suggested that this might be

due to more women having their own companions in these settings or to influence of the hospital environment and routine use of practices such as electronic fetal monitoring and epidural analgesia. It was clear that support worked in different ways in different situations (Hodnett et al. 2004).

Hodnett later developed a large RCT—the SCIL ("Supportive Care in Labour") trial—to test the effects of increasing nurse support on labour and birth (Hodnett et al. 2002a,b). The earlier trials had demonstrated that continuous support could be highly effective, but the importance of who gave the support, how support was given, and in what environment, was not clear. Observational "work-sampling" studies had shown that North American nurses only spent a small proportion of their time (6% to 24%) giving direct supportive care (McNiven et al. 1992; Gagnon & Waghorn 1996). For the SCIL trial, nurses were given a thorough induction in providing hands-on support during labour, and women were randomly assigned to receive usual or high-support care. No differences in outcome were found between the groups (Hodnett et al. 2002a). The nurses' confidence in their ability to give support was high, but their survey comments indicated that environmental factors might negatively influence its effectiveness in practice (Davies & Hodnett 2002). Drawing on the results of this study, alongside the earlier trials and her qualitative study of four "best practice" units in Canada (Ontario Women's Health Council 2002), she concluded that "the hospital culture is a powerful predictor of the likelihood of a normal, physiologic birth" (Hodnett 2002a).

In commenting on this study, and her previous work—a series of large RCTs to test different aspects of labour support—Hodnett described her thinking as having come “full circle” returning her to the point at which she began her research career, with her observations that women often went “out of labour” on admission to hospital. She argued that “researchers who seek to ask ‘Can it work?’ should be very careful about the choice of settings for their studies” (Hodnett 2002b). Ironically, however, she reported that policy makers in North America were using her trial findings to justify reduction of nurse staffing. This work has been enormously important and influential, yet she felt that it had not effectively addressed the research questions that triggered her research career. These reflections, based on her personal experience of clinical work and research and observation of birth environments, were subsequently supported by secondary analysis of trials included in her systematic review of trials of continuous labour support. Using sub-group analysis for several proxy indicators of birth environments, this review found consistent patterns suggesting that “the effectiveness of continuous intrapartum support may be enhanced or reduced by policies in the birth setting, type of provider, and timing of onset of support” (Hodnett et al. 2004).

Discussion and concluding points

The cases discussed here are not merely illustrations of problems in trial design that can be “fixed” by simply improving their planning, design, and conduct. All were considered to be well-conducted, major trials, but only the first case was really effective in addressing the research questions it set out

to answer. The dominance of the RCT in the ways in which research is reviewed, evaluated, and funded means that expensive and long-term research is being conducted that may not enhance knowledge, or may even give misleading findings if interpreted without reference to other sources of knowledge. Both the macro context of research such as political and cultural issues and the local context of the research activity are important but are not currently catered for in RCT design, which rests on positivist epistemology. This reflects a narrow concept of science, in which considerable areas of evidence are excluded, so that EBM then needs to deal with the “problem” of practice and arguments that arise about the “art” versus the “science” of treatment and care (Dopson et al. 2003).

Systems theory may offer an alternative framework for the evaluation of complex interventions and organisational changes in health care. An important tenet of systems theory, and of complexity theory, is that systems have emergent properties: The system as a whole has properties that are not present within each of its parts (Ackoff 1980). The implications of this for health research are that simplicity and control may not only be difficult to achieve in practice, but fundamentally misleading. The machine metaphor for the health system, where bodies or organisations can be analytically broken down into constituent parts to aid understanding (reductionism) then gives way to an alternative metaphor of pattern and relationship (Bateson 1985). Much of systems theory was developed in relation to organisational research, but it drew heavily on developments in mathematics and computer technology, which shifted thinking from mechanical metaphors to

informational ones. Systems theory also drew on complexity theory, which has roots as divergent as quantum physics and anthropology. Both disciplines, from very different methodological perspectives and data sources, advanced the post-modern view that truth is complex and that reality may be observer relative. The concept of relationship or embeddedness (sometimes referred to as an ecological rather than mechanical model) may be of value in developing a more critical and complex framework for evaluation of health care, including approaches to experimental research. A comparable approach, focusing on relatedness and the indeterminate nature of what constitutes evidence, is described by Wood and colleagues (1998), drawing on the philosophical work of Derrida, Deleuze, and Latour. It is reflected in the approach of Stapleton and colleagues (2002) in the informed choice study. It is also found in the increasing interest in the application of Bayesian statistics to complex change situations, from modelling fish stocks to reforms in health care (Lilford & Braunholtz 1996). Similarly, a recent article has advocated the incorporation of complex systems theory into trial design for complex interventions by moving away from the goal of standardisation to identifying core elements of a system and allowing context level adaption (Hawe et al. 2004). This responds to the principle of sensitive dependence on initial conditions found in complexity theory. Such an approach does not abandon aims of looking for general patterns but recognises that these are sought in systems where the very complexity of issues can create the impression of chaos and unpredictability.

Second, the concept of a cyclical rather than linear model of science may be of value in moving beyond the hierarchical concepts of evidence still present in EBM. In a cyclical model, knowledge development proceeds through a range of strategies and involves induction, deduction, and abduction (Downe & McCourt 2004). Experimental research designed to test or to falsify hypotheses is then understood as an important part of the scientific process, but not a superior part, or indeed the only valid form of evidence, as though it too could be isolated from its context.

My arguments, theoretically and based on the experience of this work (and indeed, many other examples of studies that could have been cited), are that RCTs can have enormous value—because they can put received wisdoms and entrenched ways of intervening in health to the test—but only if they are used appropriately and with care. RCTs should not universally be regarded as a gold standard for health research. Ideally, health research should seek to incorporate a range of methods as appropriate to the research questions. This is not simply a reiteration of the argument within EBM that trials are the gold standard for research on “treatments” (Sackett et al. 1996) but that other sources of evidence can be drawn on in other areas—such as patient preference. This argument does not challenge the linear, hierarchical approach that has become dominant in practice. When RCTs are conducted, they need to be built upon, and continue to interact with, a considerable body of prior knowledge and enquiry and a critical appreciation of the complexity of context and relationship—in short, a more critical and embedded approach.

The reductionist, linear approach on which the current concept of the trial rests reflects the epistemology of positivist science, one that is increasingly challenged by complexity theory, and by science that investigates the importance of relationship. While it may be attractive, as Green and colleagues (1998) have advocated, to try to break complex interventions, activities, or organisational forms down into their different components in order to evaluate their effectiveness more precisely, perhaps in series of trials, such an approach is in danger of splitting up the integrity of relationships upon which the potential value of such interventions depend. Instead, I suggest that further work is warranted to develop further the design of complex, non-linear trials and the integration, rather than ranking or opposition, of different research designs and traditions.

Implications for practice and research

- Evidence-based practice must be rooted in critical awareness of issues and debates regarding evidence.
- Evidence-based protocols need to be developed and used with great caution, taking account of complexity and the limitations of knowledge. They should not be applied in a uniform fashion that disregards this complexity, and the nature of probability.
- A linear hierarchy of evidence is not appropriate for many health care situations, particularly those relevant to nursing and midwifery, and should give way to a more cyclical view of knowledge development and status of evidence.
- When considering the design of research and its applications, greater regard needs to be given of context and contingency, in the complex

world of health care, which is not always amenable to a reductive and highly controlled approach.

References

Ackoff R.L. (1980). The systems revolution. In M. Lockett, & R. Spear, *Organisations as systems*. Open University Press, Milton Keynes.

Arney W. (1982). *Power and the profession of obstetrics*. Chicago: University of Chicago Press.

Audit Commission (1997). First class delivery. Improving maternity services in England and Wales. London: Author.

Barker E. (1989). *New religious movements*. London: HMSO.

Bateson G. (1985). *Mind and nature. A necessary unity*. London: Fontana.

Bero L., Grilli R., Grimshaw J., Harvey E., Oxman A., & Thomson M. (1998). Closing the gap between research and practice: An overview of systematic reviews of interventions to promote the implementation of research findings. *British Medical Journal*, *317*(7156), 465-468.

Black N. (2001). Evidence based policy: Proceed with care. *British Medical Jouornal, 323*(7307), 275-279.

Bloch M. (1989). From cognition to ideology. In M. Bloch, *Ritual, history and power: Selected papers in anthropology*. London School of Economics Monographs on Social Anthropology No 58. London: The Athlone Press.

Bourdieu P. (1972). *Outline of a theory of practice*. Cambridge, UK: Cambridge University Press.

Capra F. (1983). *The turning point: Science, society, and the rising culture*. London: Fontana.

Clifford J., & Marcus G.E. (1986). *Writing culture: The poetics and politics of ethnography*. Berkeley, CA: University of California Press.

Davies B.L., & Hodnett E. (2002). Labor support: Nurses' self-efficacy and views about factors influencing implementation. *Journal of Obstetric, Gynecologic, & Neonatal Nursing, 31*(1), 48-56.

Department of Health (2001-02). *UK statistical bulletin NHS maternity statistics* [online]. England. Retrieved May 9, 2003, from www.doh.gov.uk/public/sb0309.htm.

Dopson S., Locock L., Gabbay J., Ferlie E., & Fitzgerald L. (2003). Evidence-based medicine and the implementation gap. *Health, 7*(3), 311-330.

Douglas M. (1966). *Purity and danger: An analysis of concepts of pollution and taboo.* London: Routledge and Kegan Paul.

Downe S., & McCourt C. (2004). From being to becoming: Reconstructing childbirth knowledges. In S. Downe (Ed.), *Normal childbirth. Evidence and debate*. Oxford, UK: Elsevier.

Edwards S., Lilford R., Braunholtz D., Jackson J., Hewison J., & Thornton J. (1998). Ethical issues in the design and conduct of randomised controlled trials. *Health Technology Assessment, 2*(15), i-vi, 1-132.

Foucault M. (1973). *The birth of the clinic: An archaeology of medical perception.* London: Tavistock.

Gagnon A., & Waghorn K. (1996). Supportive care by maternity nurses: A work sampling study in an intrapartum unit. *Birth, 23*(1), 1-6.

Glasziou P., Vandenbroucke J., & Chalmers I. (2004). Assessing the quality of research. *British Medical Journal, 328*(7430), 39-41.

Goodman N. (1999). Who will challenge evidence-based medicine. *Journal of the Royal College of Physicians of London, 33*(3), 249-251.

Graham I., & Davies C. (2005). Episiotomy: The unkindest cut that persists. In C. Henderson, & D. Bick, *Perineal management: An International issue*. Massachusetts: Quay Books.

Green J., Curtis P., Price H., & Renfrew M. (1998). *Continuing to care: The organisation of midwifery services in the UK: A structured review of the evidence.* Cheshire, UK: Hochland & Hochland Ltd, Hale.

Greenhalgh T. (1999). Narrative based medicine in an evidence based world. *British Medical Journal, 318*(7179), 323-325.

Hawe H., Shiell A., & Riley T. (2004). Complex interventions: How “out of control” can a randomized controlled trial be? *British Medical Journal, 328*(7455), 1561-1563.

Hodnett E.D., Lowe N.K., Hannah M.E., Willan A.R., Stevens B., Weston J.A., et al. (2002a). Effectiveness of nurses as providers of birth labor support in North American hospitals: A randomized controlled trial. *Journal of the American Medical Association, 288*(11), 1373–1381.

Hodnett E. (2002b). Is the hospital culture a major risk factor for abnormal labour and birth? The research evidence. In *First normal birth research conference*, Final report. Lancashire, UK: University of Central Lancashire.

Hodnett E.D., & Fredericks S. (2003). Support during pregnancy for women at increased risk of low birthweight babies (Cochrane Review). In The Cochrane Library, Issue 3. Oxford: Update Software.

Hodnett E.D., Gates S., Hofmeyr G.J., & Sakala C. (2004). Continuous support for women during childbirth (Cochrane Review). In The Cochrane Library, Issue 1, 2004. Chichester, UK: John Wiley & Sons, Ltd.

Klaus M., Kennell J., Robertson S., & Sosa R. (1986). Effects of social support during parturition on maternal and infant morbidity. *British Medical Journal, 293*(6547), 585-587.

Kuhn T.S. (1970). *The structure of scientific revolutions*. London, UK: University of Chicago Press.

Langer A., Farnot U., Garcia C., Barros F., Victora, C., Belizan J., & Villar J. (1996). The Latin-American trial of psychosocial support during pregnancy: Effects on mother's wellbeing and satisfaction. *Social Science & Medicine, 42*(11), 1589-1597.

Lilford R.J., & Braunholtz D. (1996). The statistical basis of public policy: A paradigm shift is overdue. *British Medical Journal, 313*(7057), 603-607.

Mander R. (2001). *Supportive care and midwifery.* Oxford: Blackwell.

McCourt C. (2003). Social support. In C. Squire (Ed.), *The social context of childbirth*. Oxford: Radcliffe Medical Press.

McNiven P., Hodnett E., & O'Brien-Pallas L.L. (1992). Supporting women in labor: A work sampling study of the activities of labor and delivery nurses. *Birth, 19*(1), 3-8.

Medical Research Council (UK) (2000). A framework for development and evaluation of RCTs for complex interventions to improve health. Retrieved February 24, 2005, from www.mrc.ac.uk/index/publications/publications-ethics_and_best_practice/publications-clinical_trials_guidelines.htm

Norbeck J.S., DeJoseph J.F., & Smith R.T. (1996). A randomized trial of an empirically-derived social support intervention to prevent low birthweight among African American women. *Social Science & Medicine, 43*(6), 947-954.

Oakley A. (1980). *Women confined: Towards a sociology of childbirth*. Oxford: Martin Robertson.

Oakley A. (1986). *From here to maternity.* Harmondsworth, London: Penguin Books.

Oakley A. (1992). *Social support and motherhood. The natural history of a research project.* Oxford: Blackwell.

Oakley, A. (1998). Experimentation in social science: The case of health promotion. *Social Sciences in Health, 4*(2), 73-88.

Oakley A., Hickey D., Rajan L., & Rigby A. (1996). Social support in pregnancy: Does it have long term effects? *Journal of Reproductive and Infant Psychology, 1996*(14), 7-22.

Oakley A., Rajan L., & Grant A. (1990). Social support and pregnancy outcome: Report of a randomised controlled trial. *British Journal of Obstetrics and Gynaecology, 97*(2), 155-162.

O'Cathain A., Walters S.J., Nicholl J.P., Thomas K.J., & Kirkham M. (2002). Use of evidence based leaflets to promote informed choice in maternity care: Randomised controlled trial in everyday practice. *British Medical Journal, 324*(7338), 643.

Ontario Women's Health Council. (2002). Attaining and maintaining best practices in the use of caesarean sections: An analysis of four Ontario hospitals. Retrieved July 29, 2002, from www.womenshealthcouncil.on.ca/scripts/index_.asp?action=31&P_ID=1475&N_ID=1&PT_ID=13&U_ID=0.

Pieters T. (1998). Marketing medicines through randomised controlled trials: The case of interferon. *British Medical Journal, 317*(7167), 1231-1233.

Plsek P., & Greenhalgh T. (2001). Complexity science: The challenge of complexity in health care. *British Medical Journal, 323*(7313), 625-628.

Pope C. (2003). Resisting evidence: The study of evidence-based medicine as a contemporary social movement. *Health, 7*(3), 267-282.

Popper K. (1959). *The logic of scientific discovery*. New York: Harper and Row.

Ross S., Grant A., Counsell C., Gillespie W., Russell I., & Prescott R. (1999). Barriers to participation in randomised controlled trials: A systematic review. *Journal of Clinical Epidemiology, 52*(12), 1143-1156.

Sackett D.L., Rosenberg W.M.C., Gray J.A.M., Haynes R.B., & Richardson W.S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal, 312*(7023), 71-72.

Scheper-Hughes N., & Lock M. (1987). The mindful body. A prolegomenon to future work in medical anthropology. *Medical Anthropology Quarterly*, *1*, 6-41.

Sleep J., Grant A., Garcia J., Elbourne D., Spencer J.A.D., & Chalmers I. (1984). West Berkshire perineal management trial. *British Medical Journal, 289*(6445), 587-590.

Sosa R., Kennell J., Klaus M., Robertson S., & Urrutia J. (1980). The effect of a supportive companion on perinatal problems, length of labor, and mother-infant interaction. *New England Journal of Medicine, 303*(11), 597-600.

Stapleton H., Kirkham M., & Thomas G. (2002). Qualitative study of evidence based leaflets in maternity care. *British Medical Journal*, *324*(7338), 639-643.

Thomas H. (1992). Time and the cervix. In R. Frankenberg (Ed.), *Time, health and medicine.* London: Sage.

Traynor M. (2002). The oil crisis, risk and evidence-based practice. *Nursing Inquiry, 9*(3), 162-169.

Traynor M. (2003). Introduction to special issue on evidence-based medicine. *Health*, *7*(3), 265-266.

Worsley P. (1957). *The trumpet shall sound: A study of "Cargo" cults in Melanesia*. London: MacGibbon & Kee.

Wood M., Ferlie E., & Fitzgerald L. (1998). Achieving clinical behaviour change: A case of becoming indeterminate. *Social Science & Medicine, 47*(11), 1729-1738.