1. *The Nature and Use of Evidence in Midwifery Care*

Jane Munro and Helen Spiby

Introduction

At the beginning of the evidence based practice movement, much of the midwifery profession responded enthusiastically to the potential for change. Critical to this was the publication of resources of a quality not previously available to midwives, particularly *Effective Care in Pregnancy and Childbirth* (Enkin *et al*. 1989). Evidence based practice was seen to be offering a powerful tool to question and examine obstetric-led models of care that had dominated the previous decades (Page 1996; Renfrew 1997; Wickham 2000; Munro and Spiby 2001; Brucker and Schwarz 2002; Bogdan-Lovis and Sousa 2006). The results of such examination could have meant 'starting stopping' the unhelpful interventions that had embedded themselves in common practice (Muir Gray 1997). Page (1996, p. 192) even suggested that it offered to 'take us out of the dark ages and into the age of enlightenment' by demanding that women were only offered care and treatments that had been evaluated. Midwives were also becoming more active in research – undertaking studies that were to have clear clinical impact (Sleep and Grant 1987; Hundley *et al*. 1994; McCandlish *et al*. 1998). However, some midwives have not been so enthusiastic, viewing the drive to create and implement evidence as a threat to their clinical freedom (Page 1996). Bogdan-Lovis and Sousa (2006), observing the professional conflict between an obstetric and midwifery model of care, comment on the fact that in the context of over-medicalisation of childbirth, high-profile evidence is usually measuring action rather than inaction, by focusing on when to intervene rather than whether to intervene at all. They suggest that evidence based practice can thus conflict with the midwifery mandate of non-intervention in the process of normal childbirth.

What is evidence?

There continues to be considerable ambiguity about the meaning of the term *evidence* (Walsh 1996; Stewart 2001; Lomas *et al*. 2005). Lomas *et al*. (2005, p. 1)

define the concept of evidence at a basic level as 'facts (actual or asserted) intended for use in support of a conclusion'. In health care, evidence has generally been understood to be 'scientific evidence of effectiveness', which is the result of 'rigorous, objective, scientific enquiry' (DH 1996). Evidence based practice was originally defined in the medical world as 'the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients' (Sackett *et al*. 1996, p. 71). This view of 'best' evidence is also generally placed in a hierarchy. Guyatt *et al*. (2000) offered the broad definition that 'any empirical observations about the relation between events constitutes potential evidence'. Muir Gray (1997) suggested that epidemiology, the study of groups of patients and populations, was the science of most relevance to decision-making in health care.

A more inclusive definition of evidence, with a clear focus on context and implementation, was offered by the Strategic Policy Making Team (SPMT) (1999, p. 33) as 'high quality information, derived from a variety of sources – expert knowledge; existing domestic and international research; existing statistics; stakeholder consultation; evaluation of previous policies; new research, if appropriate or secondary sources'.

Lomas *et al*. (2005) undertook a systematic review to examine in detail how the concept of evidence is treated in health care by those who produce evidence, those who produce guidelines and those who make decisions. They suggest that evidence can be considered as being either colloquial or scientific. Colloquial definitions used generally in the public domain outside the research community are usually similar to 'something that points to, reveals or suggests something' (Brookes *et al*. 2004). The scientific definition, used by researchers, describes 'knowledge that is explicit (codified and propositional), systematic (uses transparent and explicit methods for methods for codifying) and replicable (using the same methods with the same samples will lead to the same results)' (Lomas *et al*. 2005, p. 3)

They found that the scientific view on evidence then breaks down roughly between two opposing views:

- that there are discoverable universal truths, independent of context;
- that evidence is of little value unless it is adapted to the relevant context.

Context-free evidence investigates what might work in ideal circumstances, and context-sensitive evidence investigates how and whether it might work in specific circumstances. Methods for obtaining evidence for either purpose are clearly very different, but as Lomas *et al*. point out, it is important that context evidence should not be viewed as any less 'scientific'. They advocate moving forward from the epistemological argument about what is 'best evidence' towards a 'balanced consensus' that is able to integrate

- medically oriented effectiveness research;
- social science–orientated research;
- colloquial evidence, representing the knowledge and views of stakeholders.

Is there such a thing as widely acceptable evidence?

Accepted knowledge is usually attached to authority and power (Foucault 1973; Oakley 1992). This dominant position can make questioning seem difficult and possibly inappropriate in what can manifest itself as a 'natural order' of status in the medical world. Downe and McCourt (2004, p. 5) describe an authoritative scientific paradigm existing in the western world that is confident that 'certain' knowledge can be established from the findings of large clinical trials and that this knowledge should be 'applied wholesale to individuals'.

The term *evidence based* is in common usage, with a confident assertion of authority (Walsh 1996; Petticrew and Roberts 2002). Lambert *et al*. (2006) identify evidence based medicine (EBM) in several different contexts: as a movement, a practice, a paradigm, a methodology, an innovation and a regulatory system. Goldenberg (2006) places 'evidence based practice' clearly in the social context of medical practice, where there is powerful established medical authority and argues that while EBM may question the practice of individual physicians, it can also reinforce the power of the medical profession as a whole, through assumptions that there is only one objective method of 'knowledge gathering'. She goes on to point out that appealing to the authority of evidence can work to obscure the subjectivity of a chosen methodology and present the evidence as 'value-free' fact rather than as the product of complex interpretation. Armstrong (2002) explores the role of EBM in supporting the collective autonomy of the 'knowledgeable' professional body but also suggests that it can overtly challenge the clinical discretion of the individual practitioner who is then expected to practice within the prescribed recommendations. In this context, there has been much resistance from the medical profession whose traditional authority has been questioned by the EBM movement that demands that 'they take science seriously' (Smith and Pell 2003; De Vries and Lemmens 2005). Although many social scientists are enthusiastic about the critique of traditional 'anecdotal' medical practice, they also articulate concerns about the objective nature of EBM (Lambert 2006). De Vries and Lemmens (2005) suggest that the cultural assumptions visible in clinical practice can also impact the collection and interpretation of evidence, and they examine the potential for financial bias, when sponsors are able to influence research design and publication.

Systems of health care often appear concerned with pathology rather than well-being and this continues to be reflected in the maternity services research, where most outcome measures are related to morbidity (Downe and McCourt 2004; National Institute for Health and Clinical Excellence 2007). Outcome measures of mortality and morbidity have an inherent authority and are key to reflecting on and developing practice. Intervention rates are also used as measures of concern. This form of evidence often guides practice by assessing the effectiveness of midwifery interventions. Downe and McCourt (2004) advocate a new framework for understanding women's experience of birth, linking maternity care and research to the promotion and exploration of positive well-being ('salutogenesis') rather than the identification and treatment of pathology.

There are several 'unscientific' sources of evidence, which are valued highly by midwives – intuition, choice, experience, insight, common sense, philosophy, policy and practice (Wickham 1999). Wickham supports the concept of 'evidence informed' rather than 'evidenced based' midwifery and describes midwifery as being 'far more than evidence', with a need to move away from the 'just science' paradigm, when recognising that only about 12% of midwifery and birth decisions can be supported by evidence (Page 1996). She goes on to suggest that midwives should not continue to look for 'absolute' answers to suit every woman, but explore how they can help women to find options that will work for them.

The methodological divide

Stewart (2001) found that definitions of evidence can vary widely among health professionals and are affected by individuals' own beliefs, which then have an impact on the perceived legitimacy of different types of evidence. Evidence that reinforced notions of authoritative knowledge appeared to move easily into clinical practice, whereas evidence that challenged professional cultural norms was more difficult to implement. Evidence that conflicts with an individual's own philosophy may thus stimulate an in-depth search for its flaws. Gergen and Gergen (2003) suggest that knowledge is generated and accepted within communities that have a shared purpose. This can mean that a research method is perceived as accurate only in terms of the conventions shared within the community. Every method of enquiry thus embodies assumptions about the nature of the world and inherently constrains ways of understanding. Experimentalists are looking at the world through cause and effect and phenomenologists are looking at the world through individuals' feelings and perceptions.

Hierarchies of evidence

Hierarchies that classify 'people or things in order of rank or importance' (Brookes *et al*. 2004) and place one group in a dominant position are a familiar and accepted concept in most areas in health care. Most of us are well aware of the different grades of practitioners and patients and are used to working with such authoritarian classifications. Different grades are a common assessment of status that pervades daily conversation in the UK National Health Service (NHS) work place.

The use of hierarchies of evidence is considered by some to be essential in order to make a distinction between evidence based and consensus based recommendations for practice (Grilli *et al*. 2000). Such hierarchies are seen by others to be limited to the underpinning of a medical positivist approach to research that places highest value on the use of quantitative methods to test hypotheses (Stewart 2001; McCourt 2005). Downe and McCourt (2004) argue that the definition of science here needs to be reclaimed and broadened in order to incorporate a fuller body of knowledge about childbirth.

A hierarchy of evidence is simply a grading system where the levels of hierarchy reflect the study design. The first attempt at grading the level of evidence was undertaken in 1979 by the Canadian Task Force on the periodic health examination (Atkins *et al*. 2004). Randomised controlled trials (RCTs) were classified as good (level 1) evidence, cohort and case control studies were classified as fair (level 2) evidence and expert opinion was classified as poor (level 3) evidence. It was very simple and as such easy to understand and implement, but criticised for its simplicity that allowed for many implicit judgements.

An example of some of the systems in current use is shown below.

The Scottish Intercollegiate Guidelines Network (SIGN) was established in 1993 to develop evidence based guidelines for the NHS in Scotland. Their experience in guideline development led them to review and update their grading systems to seek a balance between methodological rigour and applicability. They discuss the subjective nature of recommendations, which requires the guideline development group to exercise judgement based on clinical experience as well as knowledge of the evidence and methodologies used. The guideline development groups are therefore asked to 'consider the evidence in terms of quantity, quality and consistency; applicability; generalisability; and clinical impact' (Harbour and Miller 2001, p. 336). Recognising that the use of subjective judgement here can introduce bias, they defend this by stating that it is not an individual's judgement but that of a 'carefully composed multidisciplinary group'.

The Scottish Intercollegiate Guidelines Network (2008) system outlined below (Table 1.1) also discusses the recommendation of the *good practice point* (GPP) offered when there is no research evidence but when there is 'sound clinical practice that nobody is likely to question' and emphasises that it should be used only when there is no alternative. The GPP is used widely but is also an issue of some contention – when the definition of the expertise necessary to recognise such clinical practice is unclear (Miller and Petrie 2000; Walsh 2007).

The GRADE (Grading of Recommendations, Assessment, Development and Evaluation) Working Group (Schünemann *et al*. 2006) through a series of international workshops developed the following set of criteria to assess the quality of evidence (Table 1.2) and the strength of recommendations (Table 1.3). The system has explicit definitions and judgements to be used during the grading process. As they suggest, a serious limitation of this system is its complexity.

A very familiar and authoritative grading of evidence is that used by the National Institute for Clinical Excellence (NICE). NICE was established in 1999 as part of a raft of changes that aimed to reduce variations in care and maximise the cost-effectiveness of treatments within the NHS in England and Wales contained in the new policy *The New NHS: Modern and Dependable* (DH 1997) and the supplementary documentation 'A First Class Service: Quality in the new NHS' (NHS Executive 1999). The institute subsequently also incorporated the work of the NHS Health Development Agency that focused on public health and health-promoting interventions under the revised title of the National Institute for Health and Clinical Excellence. NICE is charged with the development of national guidance and its work streams include cancer service guidance, clinical guidelines, interventional procedures, public health intervention and programme

Table 1.1 SIGN grading system

Levels of evidence	
1++	High-quality meta-analyses, systematic reviews of RCTs or RCTs with a very low risk of bias
1+	Well-conducted meta-analyses, systematic reviews of RCTs or RCTs with a low risk of bias
1−	Meta-analyses, systematic reviews of RCTs or RCTs with a high risk of bias
2++	High-quality systematic reviews of case–control or cohort studies. High quality case–control or cohort studies with a very low risk of confounding, bias or chance and a high probability that the relationship is causal
2+	Well-conducted case–control or cohort studies with a low risk of confounding, bias or chance and a moderate probability that the relationship is causal
2−	Case–control or cohort studies with a high risk of confounding, bias or chance and a significant risk that the relationship is not causal
3	Non-analytic studies, e.g. case reports, case series
4	Expert opinion
Grades of recommendation	
A	At least one meta-analysis, systematic review or RCT rated as 1++, and directly applicable to the target population or,
	A systematic review of RCTs or a body of evidence principally consisting of studies rated as 1+, directly applicable to the target population and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population and demonstrating overall consistency of results; or
	Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or
	Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or
	Extrapolated evidence from studies rated as 2+

Source: Reproduced with permission from the Scottish Intercollegiate Guidelines Network (2008).

Table 1.2 GRADE system

GRADE quality assessment criteria			
Quality of evidence	**Study design**	**Lower if**	**Higher if**
High	Randomised trial	**Study quality:** −1 Serious limitations −2 Very serious limitations −1 Important *inconsistency* **Directness:** −1 Some uncertainty −2 Major uncertainty −1 *Sparse data* −1 High probability of *reporting bias*	**Strong association:** +1 Strong, no plausible confounders, consistent and direct evidence* +2 Very strong, no major threats to validity and direct evidence** +1 Evidence of a *dose response* gradient +1 All *plausible confounders* would have reduced the effect
Moderate			
Low	Observational study		
Very low			

1 = move up or down one grade (e.g. from high to intermediate); 2 = move up or down two grades (e.g. from high to low).
* A statistically significant relative risk of >2 (<0.5) based on consistent evidence from two or more observational studies, with no plausible confounders.
** A statistically significant relative risk of >5 (<0.2) based on direct evidence with no major threats to validity.

Source: Schünemann *et al*. (2006). Reproduced with permission.

guidance and technology appraisals. More recently, it has also developed tools to assist in the implementation of its guidance. NICE has published guidelines using different systems of assigning levels to the evidence and is now joining the international debate about what system is most appropriate (National Institute for Health and Clinical Excellence 2007). The one shown in Table 1.4 is that used in the Guideline for Antenatal Care (NICE 2003). It is interesting to note that guidelines published under the auspices of NICE and the collaborating centres have consistently used a grading approach but the recent Intrapartum Care Guideline does not.

As Earl-Slater (2001) pointed out, it is unclear in such hierarchies what is meant by 'well-designed' and not so 'well-designed' studies and distinguishing between the two is not straightforward. This is further complicated by the fact that there are different criteria used for judging quality and that designs are not always well documented.

Table 1.3 GRADE system

Decisions about the strength of a recommendation	
Factors that can weaken the strength of a recommendation	**Explanation**
Lower quality evidence	Will create greater uncertainty about the size of the (relative) effects (benefits and harms)
Uncertainty about the balance of benefits versus harms and burdens	Uncertainty about the baseline risk, prevalence of a problem or health status, which could affect the size of the (absolute) effects
Uncertainty or differences in values	Uncertainty about the relative importance of the benefits and downsides to those affected or differences in how important they are to different people, which could affect the balance between the benefits versus harms and burden
Marginal net benefits or downsides	The anticipated net benefits or downsides are small (and uncertain)
Uncertainty about whether the net benefits are worth the costs	Uncertainty related to lack of information about the cost or whether the resource expenditure is justified by the anticipated benefit

Source: Reproduced with permission from Schünemann, Fretheim, and Oxman (2006). *Health Research Policy and Systems* 4: 4–5, 21.

A study by Atkins *et al*. (2004) used 12 assessors who all had experience of at least one system to appraise the current six prominent systems of grading evidence. There was poor agreement among the assessors about the scoring of the six hierarchies, but consensus that none of the approaches adequately addressed all important concepts and dimensions. Their discussions, however, did agree on some conclusions including the following:

- Systematic reviews should not be included in a hierarchy of evidence because a well-done review might include anything from no studies, to poor-quality studies with inconsistent results to high-quality studies with consistent results.
- Baseline risk should be taken into consideration in defining the population to whom a recommendation applies.

These conclusions, however, are not widely endorsed.

A problem with grading evidence in this way is the fixed nature of hierarchies– with the RCT always being seen as the 'gold standard' and remaining at the top of the ladder (Petticrew and Roberts 2002). An RCT is the most reliable way of measuring the effectiveness of a treatment or intervention, as the processes employed, such as randomisation and strict inclusion criteria, minimise the risk

Table 1.4 NICE grading scheme

Recommendation grade	**Evidence**
A	Directly based on category 1 evidence
B	Directly based on • category II evidence or • extrapolated recommendation from category I evidence
C	Directly based on • category III evidence or • extrapolated recommendation from category I or II evidence
D	Directly based on • Category IV evidence or • extrapolated recommendation from category I, II or III evidence
Good practice point	The view of the Guideline Development Group
Evidence category	**Source**
Ia	Systematic review and meta-analysis of randomised controlled trials
Ib	At least one randomised controlled trial
IIa	At least one well-designed controlled study without randomisation
IIb	At least one other type of well-designed quasi-experimental study
III	Well-designed non-experimental descriptive studies, such as comparative studies, correlation studies or case studies
IV	Expert committee reports or opinions and/or clinical experience of respected authorities

Source: National Institute for Health and Clinical Excellence (2003). Reproduced with permission.

of confounding factors that influence the results (Albers 2001; Evans 2003). There are many interventions for which RCTs have not been done. They can be expensive, sometimes difficult to do and sometimes unethical as potentially harmful interventions cannot be assigned or lifesaving treatment withheld (Albers 2001). With rare events, very large samples are necessary, which can make the study not only extremely expensive but sometimes also difficult to complete. In the framework of seeking context-sensitive evidence, RCTs have significant limitations.

They cannot answer questions about patient or practitioner attitudes, beliefs and behaviours. They cannot answer questions about the prevalence and natural history of disease. They cannot answer questions about diagnostic accuracy or about reliability of examination. They cannot explore the clinical reasoning processes. For epidemiological studies relating to real-world risk factors such as smoking or the impact of caesarean sections randomisation is clearly not feasible. The need to explore processes and phenomena in detail clearly calls for qualitative studies and surveys. Information of both outcomes and processes is important to develop practice.

One of the most significant criticisms of evidence hierarchies is that they do not clearly acknowledge that research designs must be appropriate for the question (Petticrew and Roberts 2002; Goldenberg 2006). The value of the drawn-out argument about methodological primacy clearly has to be questioned (Murphy *et al*. 1998; Petticrew and Roberts 2002). Different types of research questions demand different study designs. The exclusion of evidence related to women's experiences and feelings about birth, gathered through qualitative studies, is a real problem in the context of midwifery. As Goldenberg (2006) argues, the endorsement of an evidence hierarchy that discounts evidence from qualitative research has serious implications for all areas of women's health. She points out that interventions recognising the social and political context of ill health have consistently proved to be more effective in improving health outcomes.

Commonly used hierarchies therefore privilege the RCT to the extent that this method has become the accepted paradigm for the construction of medical knowledge (Swinkels *et al*. 2002; McCourt 2005; Lambert *et al*. 2006) and offers very little to guide holistic care based on exploration of the experience. The simple outcome measures that are often used in experimental evaluations do not appear appropriate for the complex interventions involved in modern midwifery practice (Downe and McCourt 2004; Walsh 2007).

There continues to be debate about the necessity for criteria to judge qualitative research (Murphy *et al*. 1998; Sandelowski and Barroso 2002; Walsh and Downe 2005; Greenhalgh 2006; Rolfe 2006; Porter 2007). Murphy *et al*. (1998) conclude that findings from qualitative research can be readily evaluated through clear documentation of the process of data collection and analysis, in which the data are related to the circumstances of their collection. They also comment that the risk of error can be reduced where the researcher has given comprehensive attention to the analysis and reporting of negative cases. Rolfe (2006) suggests that much of this debate has simply produced a 'quality muddle' that reinforces the quantitative–qualitative dichotomy and concludes that the search for a generic framework for assessing quality should be abandoned in favour of an expectation that each individual study should be assessed on its own merits. Walsh and Downe (2005) investigated the potential for meta-synthesis of qualitative research and are enthusiastic about what it could offer to challenge the 'traditional antipathy towards qualitative evidence' (p. 210). Perhaps it is early to be confident in expectations of this technique, when the methods are not 'easy or straightforward', the possible tensions between contradictory data and findings are many and the scope of it is still being debated.

As Rychetnik *et al*. (2002) conclude, when considering criteria for evaluating evidence on public health interventions, which could describe much midwifery practice (Garrod and Byrom 2007), the term *best* evidence or *level 1* evidence is inappropriate when only one measure of evidence quality, i.e. study design, is being used. Established hierarchies cannot and do not apply to all research questions (Guyatt *et al*. 2000). It is clear that any hierarchy of evidence should be critically appraised and not 'slavishly adopted' without considering who developed it and why and the existence of evidence to support it (Earl-Slater 2002 p. 157). As Petticrew and Roberts (2002) suggested some time ago, simple typologies of evidence can be more helpful in appraising evidence for public health contexts.

What do midwives do when there is little evidence to guide practice?

Resources that have been used and have influenced midwifery practice over the years include those published by professional bodies. These can be in the form of position statements on areas of practice or topical issues, policy briefing from national reports and intercollegiate standards such as those recently set out in *Safer Childbirth* (Royal College of Obstetricians and Gynaecologists 2007). There is often a confident management reaction to such documents but implementation varies when cost implications and change management are found to be difficult.

Much of what is recognised as good front-line practice is described in the literature on clinical reasoning and defined as the 'thinking and decision-making associated with clinical practice that enables therapists to take the best-judged action for individual patients. In this sense, clinical reasoning is the means to ''wise'' action' (Jones and Rivett 2004). It is a process that includes cognition, knowledge and the ability to monitor and adjust the thinking process (Higgs and Jones 2000). There are several models of such reasoning based on analysis of practitioner and client interactions. These include hypothetico-deductive reasoning (Jones 1992; Terry and Higgs 1993), pattern recognition, and knowledge reasoning integration (Schmidt *et al*. 1990). Hypothetico-deductive reasoning, a method derived from the field of cognitive psychology, is the approach seen to dominate midwifery decision-making until the 1980s (Mok and Stevens 2005). A hypothesis is generated based on data from the woman, and then tested out or further hypotheses generated, until a care pathway is clearly defined. The hypotheses can then be confirmed by responses to the action taken; therefore, the process requires repeated reassessment.

An alternative model of clinical reasoning is based on recognition of patterns of clinical presentations. The clinician associates the present situation with other experienced cases and adopts a previously successful plan of care. Pattern recognition is thought to be possible only with a well-organised knowledge base and plentiful clinical experience; thus, it is generally only available to very experienced practitioners. Mok and Stevens (2005) suggest that experienced midwives tend to see and use patterns in a whole situation rather than reducing it to discrete

parts. The recognition of the pattern then enables them to make a judgement on the basis of a few critical pieces of information. They see this as what might be referred to as a *midwife's intuition* or *gut feeling*. In the face of unusual cases, when pattern recognition does not happen, the expert reverts to hypothesis testing. Non-expert or inexperienced clinicians tend to use the hypothesis testing model more frequently (Jones 1992). Further, clinical reasoning is an integrated skill for which a well-organised knowledge base is important but additional clinical skills are also required (Schmidt *et al*. 1990). It is the links between knowledge and these other skills which bring about effective thinking and problem solving (Alexander and Judy 1988).

Clinical reasoning in midwifery needs to incorporate both the clear and open involvement of women in decision-making and the need to collaborate with the team providing the service (Raynor *et al*. 2005). Some situations are clearly much more complex than others and the nature of the situation therefore affects the process of decision-making (Cioffi and Markham 1996). Midwives are commonly dealing with very complex and uncertain situations, where decision-making cannot be an exact science and includes many skills that include reflection and clinical reasoning (Cioffi and Markham 1996; Raynor *et al*. 2005; Hunter 2007). Cheyne *et al*. (2006) in their study examining midwives' diagnosis of onset of labour found that there were different aspects of midwives' decision-making in that they made both a diagnostic judgement and a decision about management. They used cues from their impression of the woman's appearance and other physical markers such as uterine contractions and show, as well as level of distress ranked according to perceptions of importance. They found that despite the use of these physical cues, management was not only based on this diagnosis but also powerfully influenced by the pressure from the hospital (to keep the woman at home) or the woman and her family (seeking admission).

There continues to be midwifery discussion about how to work with women in a more 'empowering way' that recognises and supports mutual interdependence (Garrod and Byrom 2007; Porter *et al*. 2007). Porter *et al*. describe a 'new professionalism' constructed in midwifery over the last decade where decision-making can theoretically take place in a negotiation between professional and client, which clearly respects each other's knowledge. However, their qualitative study into midwives' decision-making strategies concluded that midwives do not appear to have the managerial freedom to engage with women in this way, and that the decision-making remains dominated by medical and institutional authoritarianism.

Conclusion

This chapter grows from a period of work begun in 1997 of developing evidence based guidelines to support midwifery-led care in institutionalised hospital UK midwifery (Munro and Spiby 2001, 2003; Spiby and Munro 2001, 2004). This work had to engage with an ongoing struggle of having to use and justify the position of qualitative research when it was clear that understanding practice

from the woman's perspective was fundamental, and that qualitative methodologies were important in collecting this knowledge. This chapter sought to investigate and reflect on some of the issues concerned, to explore the tensions in collecting and interpreting evidence and to offer pointers for future debate and methodological development.

Walsh (2007) has recently drawn attention to a clear disappointment in the uptake of evidence based practice where evidence was originally placed in a triad of research, clinician's experience, and patient's preferences (Sackett *et al*. 1996), but where the latter two seem to have lost status in the evolved dogma. As Walsh (2007) suggests, although evidence based/informed midwifery has matured as a concept, there is a large body of evidence around normal birth that is not influencing current maternity care. We would suggest that there is now a clear place and time for acceptance of this wider body of research, which allows it to be valued and implemented if midwifery practice is going to resist further medicalisation and to develop effectively in response to women's aspirations and needs from their birth experiences. This will require clear collaboration between academics and practice based midwives working together to construct the body of knowledge. There will then need to be systems and professional leadership in place, which would be able to retain the distinctiveness of midwifery knowledge. At times, this will mean working against challenges from other professions and institutions, but will be necessary if the intention is an independent and confident authority for the profession.

References

Albers L (2001) 'Evidence' and midwifery practice. *Journal of Midwifery and Women's Health* 46: 130–136.

Alexander P and Judy J (1988) The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research* 58: 375–404.

Armstrong D (2002) Clinical autonomy, individual and collective: the problem of changing doctors' behaviour. *Social Science and Medicine* 55: 1771–1777.

Atkins D, Best D, Briss P *et al*. (2004) Grading quality of evidence and strength of recommendations. *British Medical Journal* 328(7454): 1490.

Bogdan-Lovis E and Sousa A (2006) The contextual influence of professional culture: certified nurse-midwives' knowledge of and reliance on evidence-based practice. *Social Science and Medicine* 62(11): 2681–2693.

Brookes I, Munro M, O'Donoghue E, O'Neill M, and Thomson M (eds) (2004) *Chambers Concise Dictionary*. Chambers Harrap, Edinburgh.

Brucker M and Schwarz B (2002) Fact or fiction? *International Confederation of Midwives Triennial Conference Proceedings*. Vienna, Austria.

Cheyne H, Dowding D, and Hundley V (2006) Making the diagnosis of labour: midwives' diagnostic judgement and management decisions. *Journal of Advanced Nursing* 53(6): 625–635.

Cioffi J and Markham R (1996) Clinical decision-making by midwives: managing case complexity. *Journal of Advanced Nursing* 25: 265–272.

DH (1996) *Report of the NHS Health Technology Assessment Programme*. HMSO, London.

DH (1997) *The New NHS: Modern and Dependable*. HMSO, London.
De Vries R and Lemmens T (2005) The social and cultural shaping of medical evidence: case studies from pharmaceutical research and obstetric science. *Social Science and Medicine* 62: 2694–2706.
Downe S and McCourt C (2004) From being to constructing: reconstructing childbirth knowledge. In: *Normal Childbirth Evidence and Debate* (ed Downe S). Churchill Livingstone, Edinburgh.
Earl-Slater A (2001) Critical appraisal and hierarchies of the evidence. *British Journal of Clinical Governance* 6(1): 59–63.
Earl-Slater A (2002) *The Handbook of Clinical Trials and Other Research*. Radcliffe Medical Press Ltd, Abingdon.
Enkin M, Keirse MJN, and Neilson J (1989) *Effective Care in Pregnancy and Childbirth*, Volumes 1 & 2. Oxford University Press, Oxford.
Evans D (2003) Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *Journal of Clinical Nursing* 12: 77–84.
Foucault M (1973) *The Birth of the Clinic: An Archaeology of Medical Perception*. Tavistock, London.
Garrod D and Byrom, A (2007) Public health: why the debate? *Midwives* 10(10): 470–473.
Gergen M and Gergen K (2003) *Social Construction. A Reader*. Sage Publications, London.
Goldenberg M (2006) On evidence and evidence based medicine: lessons from the philosophy of science. *Social Science and Medicine* 62(11): 2621–2632.
Greenhalgh T (2006) *How to Read a Paper*. Blackwell Publishing, Massachusetts.
Grilli R, Magrini N, Penna A *et al*. (2000) Practice guidelines developed by speciality societies: the need for a critical appraisal. *The Lancet* 355: 103–106.
Guyatt G, Haynes B, Jaeschke R *et al*. (2000) Users' guides to evidence-based medicine. *JAMA* 284(10): 1290–1296.
Harbour R and Miller J (2001) A new system for grading recommendations in evidence based guidelines. *British Medical Journal* 323: 334–336.
Higgs J and Jones M (eds) (2000) Clinical reasoning in the health professions. In: *Clinical Reasoning in the Health Professions*. Butterworth Heinemann, Oxford.
Hundley V, Cruikshank F, Lang G *et al*. (1994) Midwife managed delivery unit: a randomised controlled comparison with consultant led care. *British Medical Journal* 309: 1400–1404.
Hunter L (2007) A hermeneutic phenomenological analysis of midwives' ways of knowing during childbirth. *Midwifery*. doi:10.1016/j.midw.2007.06.001.
Jones M (1992). Clinical reasoning in manual therapy. *Physical Therapy* 72: 875–884.
Jones M and Rivett D (eds) (2004). Introduction to clinical reasoning. In: *Clinical Reasoning for Manual Therapists*. Butterworth Heinemann, Edinburgh.
Lambert H (2006) Accounting for EBM: notions of evidence in medicine. *Social Science and Medicine* 62: 2633–2645.
Lambert H, Gordon E, and Bogdan-Lovis E (2006) Gift horse or Trojan horse? Social science perspectives on evidence based health care. *Social Science and Medicine* 62: 2613–2620.
Lomas J, Culyer T, McCutcheon C *et al*. (2005) *Final Report Conceptualizing and Combining Evidence for Health System Guidance*. Canadian Health Services Research Foundation.

McCandlish R, Bowler U, van Asten H *et al*. (1998) A randomised controlled trial of care of the perineum during second stage of normal labour. *British Journal of Obstetrics and Gynaecology* 105: 1262–1272.

McCourt C (2005) Research and theory for nursing and midwifery: rethinking the nature of evidence. *Worldviews on Evidence-Based Nursing* 2(2): 75–83.

Miller J and Petrie J (2000) Development of practice guidelines. *The Lancet* 355(9190): 82–83.

Mok H and Stevens P (2005) Models of decision making. In: *Decision Making in Midwifery Practice* (eds Raynor M, Marshall J, and Sullivan A). Elsevier Churchill Livingstone, London.

Muir Gray J (1997) *Evidence-Based Healthcare How to Make Health Policy and Management Decisions*. Churchill Livingstone, Edinburgh.

Munro J and Spiby H (2001) Evidence into practice for midwifery-led care: part 2. *British Journal of Midwifery* 9: 771–774.

Munro J and Spiby H (2003) Evidence into practice for midwifery-led care: part 3. *British Journal of Midwifery* 11: 425–428.

Murphy E, Dingwall R, Greatbatch D *et al*. (1998) Qualitative research methods in health technology assessment: a review of the literature. *Health Technology Assessment*. The National Coordinating Centre for Health Technology Assessment, Southampton.

National Institute for Health and Clinical Excellence (2003) *CG 6 Antenatal Care: Routine Care for the Healthy Pregnant Woman*. NICE, London. Available from www.nice.org.uk.

National Institute for Health and Clinical Excellence (2007) *NICE Clinical Guideline 55. Intrapartum Care of Healthy Women and their Babies During Childbirth*. Available from www.nice.org.uk.

NHS Executive (1999) *Clinical Governance: Quality in the New NHS*. *Health Service Circular* 1999/065.

Oakley A (1992) *Social Support and Motherhood. The Natural History of a Research Project*. Blackwell, Oxford.

Page L (1996) The backlash against evidence based care. *Birth* 23(4): 191–192.

Petticrew M and Roberts, H (2002) Evidence, hierarchies and typologies: horses for courses. *Journal of Epidemiology and Community Health* 57: 527–529.

Porter, S (2007) Validity, trustworthiness and rigour: asserting realism in qualitative research. *Journal of Advanced Nursing* 60(1): 79–86.

Porter S, Crozier K, Sinclair M, and Kernohan W (2007) New midwifery? A qualitative analysis of midwives' decision-making strategies. *Journal of Advanced Nursing* 60(5): 525–534.

Raynor M, Marshall J, and Sullivan A (2005) *Decision Making in Midwifery Practice*. Elsevier, Churchill Livingstone, Edinburgh.

Renfrew MJ (1997) The development of evidence-based practice. *British Journal of Midwifery* 5(2): 100–104.

Rolfe G (2006) Validity, trustworthiness and rigour: quality and the idea of qualitative research. *Journal of Advanced Nursing* 53(3): 304–310.

Royal College of Obstetricians and Gynaecologists (2007) *Safer Childbirth – Minimum Standards for the Organisation and Delivery of Care in Labour*. RCOG, London.

Rychetnik L, Frommer M, Hawe P, and Shiell A (2002) Criteria for evaluating evidence on public health interventions. *Journal of Epidemiological Community Health* 56: 119–127.

Sackett D, Rosenburg W, Muir Gray J *et al*. (1996) Evidence based medicine: what it is and what it isn't. *British Medical Journal* 312: 71–72.

Sandelowski M and Barroso J (2002) Reading qualitative studies. *International Journal of Qualitative Methods* 1(1), Article 5. www.uaalberta.ca/~ijqm/.

Schmidt H, Norman G, and Boshuizen H (1990) A cognitive perspective on medical expertise: theory and implications. *Academic Medicine* 65: 611–621.

Schünemann H, Fretheim A *et al*. (2006) Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. *Health Research Policy and Systems* 4(1): 21.

Scottish Intercollegiate Guidelines Network (2008) *SIGN 50 A Guideline Developer's Handbook*. SIGN, Edinburgh.

Sleep J and Grant A (1987) West Berkshire perineal management trial: three year follow up. *British Medical Journal* 295(6601): 749–751.

Smith G and Pell J (2003) Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *British Medical Journal* 327: 1459–1461.

Spiby H and Munro J (2001) Evidence into practice for midwifery led care on the labour ward. *British Journal of Midwifery* 9: 550–552.

Spiby H and Munro J (2004) Evidence into practice for midwifery-led care: part 4. *British Journal of Midwifery* 12: 454–458.

Stewart M (2001) Whose evidence counts? An exploration of health professionals' perceptions of evidence-based practice, focusing on the maternity services. *Midwifery* 17(4): 279–288.

Strategic Policy Making Team (SPMT) (1999) *Professional Policy Making for the Twenty First Century*. Cabinet Office, London.

Swinkels A, Albarran J, Means R, Mitchell T, and Stewart M (2002) Evidence-based practice in health and social care: where are we now? *Journal of Interprofessional Care* 16(4): 335–347.

Terry W and Higgs J (1993) Educational programmes to develop clinical reasoning skills. *Australian Journal of Physiotherapy* 39: 47–51.

Walsh D (1996) Evidence-based practice: whose evidence and on what basis? *British Journal of Midwifery* 4: 454–455.

Walsh D (2007) *Evidence-Based Care for Normal Labour and Birth*. Routledge, London.

Walsh D and Downe S (2005) Meta-synthesis method for qualitative research: a literature review. *Journal of Advanced Nursing*, 50(2): 204–211.

Wickham S (1999) Evidence-informed midwifery 1: what is evidence-informed midwifery? *Midwifery Today* Autumn: 42–43.

Wickham S (2000) Evidence-informed midwifery 3: evaluating midwifery evidence. *Midwifery Today* Autumn: 45–46.