Metasynthesis: a guide to knitting smoke

Soo Downe PhD, MSc, BA, RM.
Professor, ReaCH (Research in Childbirth and Health) group, Midwifery Studies Research Unit, University of Central Lancashire, Preston, Lancashire PR1 2HE England. Email: sdowne@uclan.ac.uk

This paper was the second professor of midwifery research paper given at the Doctoral Midwifery Research Society (DMRS) on 1 November 2007 at the University of Ulster. Please visit the DMRS website: www.doctoralmidwiferysociety.org

Abstract

Since the seminal production of Noblit and Hares’ book Meta-ethnography published in 1988, the quest for an optimal route to synthesising qualitative evidence has gained momentum. Theoretical arguments in this area range diametrically from the impossibility of progressing knowledge based on qualitative evidence in the absence of such synthesis, to the impossibility of synthesising knowledge that is particular, context-specific, and dependent on the primary investigator as research instrument. Even if the need for such synthesis is accepted, there is debate at every stage of the process, from the acceptable range of inclusion of studies for reviews to the place of formal search strategies, inclusion criteria, quality assessment, and the optimum method of synthesis.

This paper explores all of these dimensions and dilemmas, with reference to the theoretical and methodological literature in this area. It also discusses solutions that have been employed by the author and her colleagues in undertaking a series of metasynthesis-based reviews in the area of maternity care.

Key words: Metasynthesis, qualitative research, theory generation, midwifery research, research dilemmas

Background

Very few researchers and practitioners would now dispute the value of combining the findings of randomised clinical trials to establish best evidence for guidelines and practice. While there is still some debate around the details of the methods that should be used, and considerable disagreement about how much weight meta-analysis should have in dictating best practice, this is in the context of the widespread authority of the techniques for systematic review and meta-analysis set out in the Cochrane handbook (Higgins et al, 2006). Criteria for assessing the quality of studies to be included are also standardised, both at the level of individual randomisation (Moher et al, 2001) and for cluster randomised trials (Campbell et al, 2004).

However, there is much less agreement around how to combine qualitative studies. Indeed, there have been fundamental arguments about whether this is a good thing to do at all. This is a live argument that has generated an increasing number of methodological and philosophical papers over the last ten years (Sandelowski et al, 1997; Silverman, 1997; Barroso and Powell, 2000; Beck, 2002; Sandelowski and Barrow, 2003; Finfgeld, 1999, 2003; Walsh and Downe, 2005, 2006; Downe et al, 2007; Finlayson and Dixon, 2008; Simkhada et al, 2008).

As Walsh and Downe note (2005), Stern and Harris (1985) first used the ‘qualitative metasynthesis’ in relation to combining the findings of a range of studies. However, it is the methods described in the seminal publication of Noblit and Hare (1988) that are cited most frequently in syntheses of qualitative studies currently. Although their technique was specifically focused on what they termed ‘meta-ethnography’, it has been extended to combinations of studies with a wide range of theoretical perspectives. This multiplicity is magnified by the addition of a range of analytic strategies, that have been typified as theory building (divided into grounded formal theory and metastudy), theory explication, and descriptive metasynthesis (Finfgeld, 2003). Against this multiplicity of approaches, there appears to be general agreement that, as Deborah Finfgeld notes: ‘The goal of metasynthesis is to produce a new and integrative interpretation of findings that is more substantive than those resulting for individual investigations’ (Finfgeld, 2003: 894).

Though the technique has a recent history, it has been applied in a range of areas, including transformational leadership (Pielstick, 1998), experience of chronic illness (Thorne et al, 2002), diabetes (Paterson, 2001; Campbell et al, 2003), concepts of caring (Sherwood, 1997) postnatal depression (Beck, 2002), adolescent motherhood (Clemmens, 2003), midwifery care (Kennedy et al, 2003), and midwifery expertise (Downe et al, 2007).

At least three active debates can be identified in the current metasynthesis literature. These cover the philosophical position taken by authors, the importance or otherwise of methodological rigour or quality in the studies to be included and the analytic approach employed. To some extent, these areas are interrelated. The next section explores the debates, and describes some of the specific approaches taken by the author and her colleagues in metasynthesis studies they have recently undertaken.
Discussion of the three debates

Philosophical positions: finding the truth or reducing uncertainty

Most qualitative research is seen as a specific construct of the world that is negotiated between the study participants, their social context, the researchers, and the reader of the research. Qualitative study tends to be based on relativist theoretical perspectives, such as interpretivism (‘the world exists, but what we perceive as ‘reality’ is only a social interpretation’) or constructivism (‘the world only exists as an artefact of human interpretation and perception’) (Crotty, 1998). In theory at least, research carried out under both conditions tends to limit any claim that the findings represent ‘the (absolute) truth’ about the way the world is. Sandelowski and her colleagues (1997) express this elegantly as they argue against ‘summing up’ the knowledge generated by this kind of research: ‘To summarise qualitative findings is to destroy the integrity of the individual projects on which such summaries are based, to thin out the desired thickness of particulars… and ultimately to lose the vitality, viscerality and vicarism of the human experiences represented in the original studies’ (Sandelowski, 1997: 366).

Apparently in direct contrast, in the same year that Sandelowski published her critique, Silverman argued that qualitative researchers risk increasing marginalisation from policy and practice if their work remains isolationist and esoteric, and if they are unwilling to engage in synthesis that might provide middle range theories to advance knowledge and practice (Silverman, 1997). Earlier, Statham had referred to this tendency to curtail the transferability of findings by refusing to move beyond single studies as ‘analytic interruptions’. In a somewhat controversial statement that extends this critique, Jensen and Allen (1996) wrote about the need to search for ‘truth value’.

Over the ten years or so since these views were expressed, acceptance of the need for increased transferability of study findings has become dominant, and metasynthesis studies are now appearing with increasing regularity, as noted above. Contemporary discussions seem to be more about method than about the fundamental philosophical principles of qualitative data synthesis. The ongoing debate has generated middle ground positions that are now gaining prominence. This is the stance taken by Kearney, for example: ‘We seek to develop methods in which alternative experiences and interpretations are revealed rather than neutralised… (in which) differences must be sought and honoured’ (Kearney, 1998: 499).

Recent observations by Noblit in a paper that presents the opinions of five of the leading proponents of a range of metasynthesis approaches suggest this was the original theoretical position that underpinned the meta-ethnography he undertook with Hare in 1988 (Thorne et al, 2004). He goes on to comment that the 1988 book was based on what he called ‘three moves’: that summaries of qualitative research offer interpretation, not new data; that what is generated are metaphors, not answers; and the notion that the socially-driven explanation that arises from such a synthesis is, in itself, a translation of findings from one context to another. Noblit saw the result of such an enterprise as contingent. It may or may not resonate with the context in which readers find themselves, so the usefulness of findings from metasynthesis studies will vary by how far they offer effective sets of metaphors for understanding in other contexts. When the derived metaphors, or ‘line of synthesis’ can be effectively translated, this results in what Louise Jenson has summarised as the reduction, but not the elimination of uncertainty (Thorne et al, 2004: 1342).

Debates around rigour

Louise Jensen probably takes the most robust and definitive approach to methodological quality in metasynthesis, warning, in another evocative turn of phrase, of the danger of creating ‘metasoup’ if the need for rigour is not appreciated (Thorne et al, 2004: 1347). As she acknowledges, her background in meta-analysis has strongly influenced this stance. In contrast, Sandelowski remains somewhat equivocal. She critiques the ‘presumptive assumption baggage that comes with the analogy to meta-analysis’ (Thorne et al, 2004: 1361), but she has moved to a position of relative acceptance of the need for agreed standards at some level: ‘We envision metasynthesis evolving into a complex set of strategies with which scholars in the health research field render the warranties for their qualitatively derived assertions in a manner that is increasingly accessible, auditable, and transparent’ (Thorne et al, 2004: 1361). This is the stance taken by the author of this paper and her colleagues (Walsh and Downe, 2005, 2006).

Scope of the review phase

The differences of opinion about what metasynthesis is or should be have also led to disagreement about the scope of metasynthesis reviews. At one extreme, Paterson and colleagues (2001) attempted to synthesise 292 qualitative studies of chronic illness, in a technique they termed ‘metastudy’. Not surprisingly perhaps, Thorne, who was a member of the team that commented that ‘truth claims of all sorts slipped into and out of focus’ (Thorne et al, 2004: 1336). At another extreme, metasynthesis has been undertaken with only three studies (Russell et al, 1997). Noblit and Hare took the line that ‘few studies are sufficient’ (Noblit and Hare, 1988), but did not define ‘few’. There is as yet, no agreement on whether search strategies for studies to include should be narrow or widely focused, or how strictly they should be applied. The approach of the author and her colleagues to this has been to be explicit about what question they started off with, and where they ended up, and to take an iterative approach to which studies to include. For example in our metasynthesis of expertise (Downe et al, 2007), we started with the question: ‘What is the nature of midwifery expertise in the context of physiological birth?’

After a series of iterations around testing the topic against the current literature and extensive debate, we ended up with a much more nuanced question: ‘What accounts of intrapartum midwifery skills, practices, beliefs and philosophies are given by practitioners working in the field of midwifery who are practising “beyond the ordinary” in the intrapartum setting?’
As Downe et al (2007) explain, this change came about as a result of significant reflection on, and reflexivity about our personal political and ideological positions, the professional project of licensed midwifery, and midwifery mythologies that may or may not be justified. Most fundamentally and controversially, we began to question our own definitions of ‘midwife’, and of expertise. This led to the change in scope for the review.

Having defined the area of investigation, Downe et al (2007) use formal search strategy methods, augmented with techniques termed ‘berry-picking’ (Bates, 1989), where initial searching against the broad topic leads to new, unpredictable ideas and directions, and even a new formulation of the original query. The search also lead to ‘backchaining’, which involves close searching through reference lists of included studies to identify any further papers that may fall within the scope of a review. We have also adopted the techniques of theoretical saturation and of searching for disconfirming data, borrowed from grounded theory. In the case of a metasynthesis, the unit of analysis for these techniques is a study. As we analysed each additional study, we consciously checked if the findings extended or refuted the emerging line of argument synthesis (Downe et al, 2007). If additional studies continue to reinforce the line of argument, it is likely that continuing to search for new studies will reap increasingly diminutive returns, and so the search can be truncated at that point. These approaches align somewhat with Noblit and Hare’s concepts of reciprocal and refutational synthesis, which are discussed further below (Noblit and Hare, 1988).

A range of other approaches to selection is evident in the literature. For example, in their study of the experience of living with diabetes, Campbell et al (2003) selected a purposive sample of ten studies, having decided a priori that ten would be sufficient for the aim of their study. Any approach can be valid, as long as the teams using them understand and pay attention to the underlying theoretical perspective adopted, and are explicit about the strengths and weaknesses involved, and the limits these impose on what they claim their findings say about the particular area of study they have approached.

Quality of included studies

One of the most active debates in the area of qualitative synthesis, and indeed, qualitative research in general is the application of quality criteria. Adherents of a strongly constructivist persuasion are most resistant to this approach, believing that any externally imposed rules of acceptability for context-specific in-depth studies risks violating epistemological principles of knowledge as particular, specific, and resistant to exact replication. At the other extreme, a wide range of checklists have been proposed, with varying degrees of rigidity. In a search for a quality tool, Walsh and Downe (2005) located eight existing checklists and summary frameworks, covering a wide range of elements. Their review included the summary framework of Spencer and colleagues (Spencer et al, 2003), which was in itself based on 29 different checklists. Recently, a team undertaking a review of access to antenatal care needed to develop another tool to assess the quality of qualitative studies in their sample (Simkha-da, 2008), which suggests that this issue is still live in the qualitative research community.

If the point of undertaking qualitative research is to change policy and practice, the pragmatic approach to assessment of quality is probably somewhere in the middle of the extreme stances described above, as Murphy and colleagues have observed: ‘Some argue that… the very idea of criteria is incompatible with… the… anti-realist assumptions… (of qualitative research)…. We suggest that this position is unnecessarily constraining… if the findings of research cannot be taken to represent even an approximation of the truth… why should commissioners… fund… such research’ (Murphy et al, 1998: 10).

In the same vein, Barbour (2001: 1115) argues for checklists to be viewed as ‘reflective rather than constitutive of good research’. Sandelowski and Barroso (2002) proposed a lengthy list of quality markers, but go on to urge that this should be used flexibly. From their position, all studies may contribute to an emerging understanding of a field, no matter how many ticks they generate on a list.

In contrast to the relativist position of Sandelowski and Barroso, Walsh and Downe (2005) have argued that ‘metasynthesis of methodologically flawed studies may result in flawed metasynthesis’. This statement is based on the belief that a study, which does not meet at least a minimum level of credibility, transferability, dependability, and confirmability (Lincoln and Guba, 1985) is unlikely to contribute to a significant reduction in uncertainty in the area under scrutiny. Walsh and Downe’s quality assessment tool was published in 2005, and since then, a grading system for the studies included was adopted, adapted from an initial idea of one of our Masters’ students (with thanks to Eileen Whitehead) (Downe et al, 2007) (see Box 1). The authors do not include studies that score less than C+ on the tool.

The tension between those who believe that any qualitative study is of value as it captures (at least partially), the essence of specific human experiences, and those who want to take a strictly quality-mediated approach is still not resolved. Again, anyone working in this area will need to decide at the outset where they are on this continuum, justify their position, and carry out their research accordingly.

Box 1. Quality summary score for qualitative studies

<table>
<thead>
<tr>
<th>Key to quality rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>A – No or few flaws: The study credibility, transferability, dependability, and confirmability is high</td>
</tr>
<tr>
<td>B – Some flaws, unlikely to affect the credibility, transferability, dependability, and/or confirmability of the study</td>
</tr>
<tr>
<td>C – Some flaws, which may affect the credibility, transferability, dependability, and/or confirmability of the study</td>
</tr>
<tr>
<td>D – Significant flaws, which are very likely to affect the credibility, transferability, dependability, and/or confirmability of the study.</td>
</tr>
</tbody>
</table>

(Downe et al, 2007: adapted from Jackson, unpublished)
Box 2. Analytic strategy from meta-ethnography: Noblit and Hare (1988)

- Reading the studies
- Determining how the studies are related
- Translating the studies into one another
  - Reciprocal translation
  - Refutational translation
- Synthesising translations: ‘line of argument’
- Expressing the synthesis.

Analytic strategy
In general, the aim of metasynthesis is to look for patterns in the included data that have high explanatory power for the phenomenon under investigation. Before this can be determined, decisions need to be made about which data are to be included, and about how patterns are to be identified.

Which data to include?
Once agreement is reached on which studies are included in a review, there is a question to be resolved about which data to include in the analysis. It might be argued that original transcripts are closest to the phenomenon under scrutiny. However, even these are only interpretations of those phenomena by the original researchers. The approach taken by Britten and colleagues (2002), where some of the included original papers had at least one member of the review team as a researcher, is one way of dealing with this factor of constructed or interpreted data. However, in general this is not a practical approach to take. The next level of interpretation is the original report of the data, or the thesis if it is a PhD, and this might be the ideal source for review groups that were not a part of the original data collection team. Often however, it is the drastically pared down account of a study that is included in a journal, which is the primary source for metasynthesis studies. The limitations this imposes need to be acknowledged by all those working in this field.

Analysis
While a range of approaches to analysis has been taken, the techniques described by Noblit and Hare in 1988 are probably the most widely used, in whole or in part (see Box 2).

At the analysis phase, they proposed three distinct stages (or ‘moves’) of analysis. The first is termed ‘reciprocal’, and entails a search for phrases, metaphors and themes that occur repeatedly across the included data. The second is termed ‘refutational’, and involves a conscious search for phrases, metaphors and themes that refute any emerging patterns. The third phase is termed the ‘line of argument synthesis’, and results in the summary statement that most completely expresses the emerging patterns across the included studies. Despite the frequency with which Noblit and Hare’s 1988 work is cited, Noblit observed in 2004 that, while most published metasynthesis accounts describe reciprocal findings, few report on the refutational phase of the work, and even fewer reach a distinct ‘line of argument’ synthesis (Thorne et al, 2004: 1349). Indeed, in critiquing the proliferation of under-theorised metasynthesis studies, Margarete Sandelowski has commented that ‘we are in an era of metamadness… or even ‘metajeopardy’ (Sandelowski, 2006: 11). Like Noblit, she notes a difference between what she calls qualitative metasummary, which is a rather quantitatively orientated aggregation of qualitative findings, and qualitative metasynthesis, which is seen as an interpretive integration of qualitative findings that are themselves interpretive syntheses of data. In the latter case, Sandelowski argues that what is intended is something that is more than the sum of its parts. Her key argument, stated forcefully, is the need for those doing metasynthesis to ‘come to a point’.

As an example of how ‘coming to a point’ may look, Downe et al’s (2007) study of maternity care practitioners who were practising beyond the ordinary reached the following line of synthesis: ‘As the practitioners in the studies in our review become more expert, they appeared to (re)value and to express qualities such as trust, belief and courage, to be more willing to act on intuitive gestalt insights, and to prioritise connected relationships over displays of technical brilliance… in some of the accounts, the enactment of vocation led these experts to move outside of and beyond normative childbirth practices, and so to become more exposed to critique’ (Downe et al, 2007: 136).

Emerging fields of enquiry in reviews of non-controlled studies
As this paper has indicted, there is an emerging consensus for some aspects of metasynthesis. However, a wide range of debates persist, as Thorne and colleagues have noted: ‘Qualitative metasynthesis seems to be unfolding before our eyes as we dive deeper and deeper into its implications and applications… the methodological conventions remain in flux…’ (Thorne et al, 2004: 1345).

Beyond these debates, there is increasing interest in what has been termed ‘realist policy review’ (Pawson and Tilley, 1997; Pawson et al, 2005). This approach tends to include studies that utilise a range of methods, and even the opinion literature. It seeks to identify ‘what works, for who, in what context’. Qualitative research is fundamental to this endeavour, and the potential for combining insights from a range of methodological approaches offers the potential for a much richer, more flexible, and dynamic approach to evidence generation, and to guideline and practice development, built firmly on the premise of limiting uncertainty, as opposed to seeking absolute truths. Downe et al (2007) have explored the potential for this way of seeing in terms of normal birth and more recently, Murray Enkin has charted his personal move away from absolutist faith in trials’ evidence towards a more relativist, contingent position: ‘This paper… was conceived during an era of medical authoritarianism, born in a time of nascent… family-centred maternity care, matured in a period of enthusiastic (but not unquestioning) homage to evidence-based obstetrics, and culminated in a reluctant but comforting acceptance of uncertainty… It is, to use an ancient word I only recently learned, a clinamen, a swerve, a point of intellectual revision…’ (Enkin et al, 2006: 265).
Conclusion
The process of undertaking metasynthesis is iterative, contingent, and never definitively complete. There is debate about most steps in the process. However, there does appear to be an emerging consensus that it is worth doing, and that the findings can add to the general sum of knowledge in specific areas, primarily by reducing uncertainty to a greater or lesser extent. While there is no definitive ‘how to’ guide in this area, this does not mean that ‘anything goes’. Each decision made in the process of undertaking such a review requires attention to one’s theoretical perspective, epistemological stance, and methodological integrity. In doing this, those aiming to make some sense of the process would do very well to heed the following words of Margarete Sandelowskii: ‘Ask not what qualitative research can do for you, ask what qualitative research can do with… you and what you can do better with qualitative research. Answer not that qualitative health research is only about asking the right questions, counter that… [it]… is about answering them too’ (Sandelowski, 2004: 1383).

References