**DETAILS OF APPROACHES TO SYNTHESIS**

**A METHODOLOGICAL APPENDIX TO THE PAPER:**

**SYSTEMATICALLY REVIEWING QUALITATIVE AND QUANTITATIVE EVIDENCE TO INFORM MANAGEMENT AND POLICY MAKING IN THE HEALTH FIELD**

NICHOLAS MAYS

Professor of Health Policy

Health Services Research Unit

Department of Public Health and Policy

London School of Hygiene and Tropical Medicine

London University

UK


CATHERINE POPE

Senior Research Fellow

School of Nursing & Midwifery

University of Southampton

Southampton

UK


JENNIE POPAY

Professor of Sociology and Public Health

Institute for Health Research

University of Lancaster

Lancaster

UK

**Introduction**

In the paper with which this appendix is linked we have identified four basic approaches to evidence review: narrative approaches (which may include research and non-research based evidence both qualitative and quantitative and typically deal with the findings of included studies in their own terms rather than transforming one form of 'data' into another); qualitative (which convert all available evidence into qualitative form using techniques such as 'meta-ethnography' and 'qualitative cross-case analysis'), quantitative (which convert all evidence into quantitative form using techniques such as 'quantitative case survey' or 'content analysis') and Bayesian meta-analysis and decision analysis (which can convert qualitative evidence such as preferences about different outcomes into quantitative form or 'weights' to use in quantitative synthesis). We recognise that the boundaries between these four 'types' of approaches is somewhat permeable: some approaches we describe as qualitative are essentially narrative in form but this typology provides a heuristic device to help us organise the material. In this appendix we provide some more detail of the approaches discussed in the paper and reference to more technical discussions where appropriate.

**1.      Narrative approaches**

There is no clearly delineated body of knowledge describing particular narrative approaches to evidence synthesis and substantive discussion of questions of rigour and reliability in relation to these approaches is not readily available. Various terms are used, including 'narrative review', and more recently, 'narrative synthesis' and 'realist synthesis'. These can mean different things to different writers In broad terms, narrative approaches summarise , compare, explain and interpret evidence of all types relevant to a particular question. In the past, these approaches tended to lack transparency and to summarise findings of included studies rather than attempting a synthesis. Typically, however, even more recent approaches do tend to deal with the findings and interpretations from published studies in their own terms, without any attempt to transform them for analytical purposes.

Dixon-Woods and colleagues[1] argue that narrative approaches are very flexible, allowing for different types of evidence – both qualitative and quantitative – to be reviewed, though not necessarily allowing full integration. This flexibility and ease of handling of a very wide range of evidence means that such approaches are likely to remain an important tool for policy and management-relevant reviews. Narrative approaches also allow the development and testing of theories and explanations that attempt to account for all the published findings relevant to a question.

## 1.1 Narrative reviews

### What are they?

The term 'narrative review' has been used to describe the traditional literature reviews (see for example, Dixon-Woods [1] and Pawson and Bellaby [2]), of the type still widely undertaken in the social sciences. In the past, these types of narrative reviews have been concerned with questions ranging well beyond whether a programme or interventions works, considering other questions such as: 'What do we know about the causes of a particular social and/or health problem? What are the implications of evidence on causality for the type of programmes or interventions that should be developed?' Increasingly, however, narrative reviews are also addressing questions of effectiveness and cost-effectiveness. There are many classic examples of narrative reviews addressing a range of questions relevant to policy and practice. The review of research about antisocial behaviour by Rutter [3] and colleagues is a good illustration of the wide-ranging nature of narrative reviews, where the authors use a writing style that helps to draw the reader through a complex array of material in a way that allows a clear (and, it is hoped accurate) picture to be painted. Often, the style of discussion used is that of telling a story or developing a line of argument, using a range of studies that each contributes to the 'plot'. Narrative reviews may include non-research sources of evidence, and may be less concerned with assessing the quality of evidence and more focused on gathering relevant information that provides both context and substance to the authors' overall argument. Indeed, the Rutter et al. review mentioned above was undertaken at a time when the methods for systematic, comprehensive literature retrieval and quality assessment were developing.

### What are the Strengths of narrative reviews?

Narrative reviews still do not usually follow formal standardised procedures (though they are tending towards greater formality and explicitness in the context of the drive for greater rigour in evidence review). However, Pawson and Bellaby [2] have argued that it would be wrong to assume that the lack of standardised methods inevitably means that there is no logic to the method. They suggest that narrative reviews focusing on whether particular programmes or interventions work are based on a 'configurational' approach to causality'.

> 'According to this perspective narrative reviews of evidence on the effectiveness of programmes or interventions are built on the assumption that positive outcomes will result from the combination of a series of program/intervention attributes....Interventions work, it is considered, because of the compatibility of target group, setting, program stratagem, program content, implementation details, stakeholder alliances and so on. When using this framework it is the entire 'recipe' that makes the difference. All of these

ingredients, along with information on outcomes and on the methodology employed in the original evaluations constitute the ontology, or the how, of narrative review.' [2].

From this perspective, the task of the reviewer is to identify studies that provide the richest description of the significant properties of a particular programme or intervention.. Narrative review can therefore identify examples of 'good practice' or 'best buys' based on a judgement of the 'fit' between an intervention or programme and the critical success factors the review has identified

### What are the limitations of narrative reviews?

There are many drawbacks to narrative reviews.  In the past it was rare for narrative reviewers to provide details about decisions regarding searching or study selection or to fully describe how the review was conducted. The review process, therefore, lacked transparency. However, narrative approaches are become more explicit in their methods. This is occurring in parallel to the emergence of a new narrative approach to evidence review more appropriately termed narrative synthesis and discussed more fully below.

There are also many practical challenges. The very flexibility of narrative reviews means that the number of studies and other sources that can potentially be included could become unmanageable as could the amount of information to be extracted from studies. Additionally, the diversity in the type of research makes the appraisal of study quality difficult, presents particular problems for the extraction of data in a common format and makes it hard to weight different types of evidence.

Perhaps the most important criticism of narrative reviews is the potential for bias and hence for unreliable conclusions to be drawn.  Shadish and colleagues[4] suggest that in order to claim generalisability narrative reviews have to demonstrate 'proximal similarity'. This process involves selecting a feasible number of studies to review rather than attempting to be comprehensive, and choosing a manageable number of programme characteristics to explore in detail from what would certainly be a much larger number.  Pawson and Bellaby[2] question the logic of 'proximal similarity', suggesting that in this  process some studies and factors will be privileged over others and that this introduces a whole range of biases - from those associated with publications to those arising from the personal orientation and interests of the reviewer.  However, the extent to which this problem is unique to narrative reviews should not be exaggerated. Recent debates in quantitative meta-analysis of trials of effectiveness have often turned on which studies to include and how much weight to give to each.

**1.**

## 1.2 Narrative synthesis

### What is it?

Like narrative review, 'narrative synthesis' involves a narrative approach to evidence review but it differs from narrative review in including a formal analytical process of synthesis to generate new insights or knowledge and by seeking to be systematic and transparent. At least two different approaches to narrative synthesis can be identified. The first of these, realistic synthesis, is discussed in more detail in a complementary report which is also available on the websites of the two funders of the current project[5]. The second, which is referred to as narrative synthesis, is described more fully below.

Guidance for the conduct of narrative synthesis, is currently being developed in a project funded by the UK Economic and Social Research Council[6]. In this guidance, narrative synthesis is described as incorporating six main 'stages' although recognising that the review process will be iterative rather than linear: (1) identifying the broad focus of the review, and searching for and mapping the available evidence; (2) specifying the review question; (3) selecting studies to include in the review; (4) data extraction and study quality appraisal; (5) the synthesis; and (6) reporting the results of the review and dissemination. The draft guidance on the conduct of narrative synthesis suggests that a narrative synthesis approach could be used in three situations: before undertaking a statistical meta-analysis; instead of a statistical meta-analysis because the experimental or quasi-experimental studies included are not sufficiently similar to allow for this; and where the review questions dictate the inclusion of a wide range of different research designs, producing qualitative and/or quantitative findings, and/or non-research evidence, for which no other specialist approach to synthesis is appropriate. The guidance provides advice on the conduct of narrative synthesis and describes some specific tools and techniques that can be used in the synthesis – these include thematic analysis discussed further below. Whilst narrative synthesis can involve the manipulation of statistical data, the defining characteristic is that it utilises a narrative interpretative approach to the process of synthesis.

It has been suggested[1] that different types of evidence synthesis can be located along a continuum from quantitative approaches, which involve the pooling of findings from multiple studies (e.g. meta-analysis), to qualitative approaches, which involve an interpretative approach (e.g. meta-ethnography). Narrative synthesis lies between these two. It will always involve the 'simple' juxtaposition of findings from the studies that have been included in the review. However, where the evidence allows, it will also involve some element of integration and/or

interpretation. Popay and colleagues identify three main elements to a narrative synthesis process[6]:

- Developing a preliminary synthesis of the findings of included studies;
- Exploring relationships in the data ;
- Assessing the robustness of the synthesis product.

As they note: 'The key purpose of the narrative synthesis remains the organisation, description, exploration, and interpretation of the study findings and the attempt to find explanations for (and moderators of) those findings'.

### *What are the strengths of narrative synthesis?*

The developing methods for narrative synthesis are addressing many of the drawbacks of narrative reviews identified above. Narrative synthesis is more transparent and systematic and provides a framework for making choices about the appropriate methods to use to synthesis data from disparate sources. The approach retains the flexibility of the narrative review being appropriate for a wide range of review questions and allowing for the inclusion of wide ranging and disparate types of evidence – research and non-research. A variety of specific methods – appropriate to the review question and the evidence to be synthesised – can be utilised within the narrative review framework as can evolving methods for the management of larger numbers of studies and for the complex process of study quality appraisal. Perhaps most importantly, this approach adopts a formal process to deal with potential biases introduced through the decisions on what research to include and exclude, and/or the synthesis process itself. Hence, it should produce more reliable and generalisable conclusions than narrative reviews.

### *What are the limitations of narrative synthesis?*

Formal guidance on the conduct of narrative synthesis is currently being developed so one limitation of this approach is that at present it has not been extensively practised and there are no fully worked through examples of the use of the approach. In some senses, it is a general framework within which a wide range of specific methods for synthesis can be used and how reviewers make decisions about the appropriate methods to use in their review is a question that remains to be explored.

### 1.3.    Thematic analysis

### *What is it?*

'Thematic analysis' comprises the identification of the main, recurrent or most important (based on the specific question being answered or the theoretical position of the reviewer) issues or

themes arising in a body of literature. It is one of the most common methods for identifying, grouping and summarising findings from included studies in narrative review and can also be used in the early stages of a narrative synthesis.. Though thematic analysis is primarily qualitative in origin, themes can be counted and tabulated (much as in 'content analysis' (see below)). However, it is perhaps more common for a thematic analysis than a content analysis to be developed in at least a partially inductive manner; i.e. without a complete set of a priori themes to guide data extraction and analysis from the outset.

Thematic analysis tends to work with, and reflect directly, the main ideas and conclusions across a body of evidence, looking for what is prominent rather than developing 'higher order', new explanations for findings that do not appear in any of the published accounts of individual studies (this contrasts with approaches such as meta-ethnography, below).

### *What are the strengths of thematic analysis?*

The advantages of thematic analysis are that it provides a means of organising and summarising the findings from a large, diverse body of research. Like narrative review, it can handle qualitative and quantitative findings since it is still, in large part, a narrative approach. This means it can be applied in almost all circumstances.

### What are the limitations of thematic analysis?

As with narrative reviews, the flexibility of thematic analysis is associated with a lack of transparency. It can be difficult for the reader to be sure how and at what stage themes were identified. Would the review have looked very different if an entirely a priori, theoretically-driven approach had been used instead of an inductive approach in which themes 'emerge' from the process of analysis? This uncertainty reflects the fact that thematic analysis can be undertaken in very many ways (i.e. quantitatively or qualitatively; inductively or deductively; theoretically driven or descriptively). It is also unclear whether the findings from thematic analysis should reflect the frequency with which each theme is reported or its explanatory significance (content analysis suffers similarly).

### 2. Qualitative approaches

### 2.1 Meta-ethnography

### *What is it?*

'Meta-ethnography'[7] is a way of using qualitative research techniques to synthesise multiple written interpretive accounts (i.e. a number of qualitative research reports). The output of a meta-ethnography is a new, 'higher order' interpretation or theory that satisfactorily accounts for the

available body of evidence. It is an approach rooted in the recognition that the accumulative logic of science (in which lots of individual studies build a larger knowledge base) is too seldom applied in the field of ethnographic qualitative research where distinctiveness is emphasised. It also grew out of a dissatisfaction with the traditional narrative review which "lacks some way to make sense of what the collection of studies is saying" [7] (pp14-5). Meta-ethnography is interpretive rather than aggregative: it seeks to do more than simply collect and review a series of accounts (as in a narrative review) and instead aims at a novel synthesis which develops theory to explain the range of research findings encountered. It is thus a way of re-analysing and comparing the texts of published studies (normally, rather than the original data of each) to produce a new interpretation of the findings. Meta-ethnography is both inductive and interpretive, using specific research methods such as 'thematic analysis' and 'constant comparison' to construct a higher order synthesis and hence new knowledge [8].

Noblit and Hare[7] are careful to differentiate this approach from methods for synthesising quantitative research (e.g. meta-analysis). The term 'meta' in this context refers not to generalisation from a group of studies, but to the translation of studies into one another. This translation is idiomatic – it focuses on the translation of salient categories of meaning, rather than the literal translation of words or phrases. The product of this process of translation is not like that of a meta-analysis (not simply pooling data) or narrative review.

The process of meta-ethnography entails seven steps:
1. Getting started: identifying an intellectual interest that qualitative research might inform;
2. Searching and selection of relevant studies: as in qualitative research this is purposive, it may not need to be exhaustive or comprehensive ;
3. Reading the studies repeatedly: to identify metaphors/concepts or explanatory schema (in effect these are the data for the synthesis), maintaining as far as possible the original terminology and remaining faithful to the original meanings;
4. Determining how the studies are related: compiling a list of the metaphors or explanatory schema, phrases, ideas and key concepts (and their relations) used in each of the studies and juxtaposing them;
5. Translating the studies into one another: comparing the metaphors/ explanatory schema, looking at similarities and interactions between them in the different studies;
6. Synthesising the translations: comparing the various translations produced, to identify types of translation, or metaphors/explanatory schema that transcend individual

accounts, and in turn can be used to produce new interpretation or conceptual development;

7. Expressing the synthesis: communicating the synthesis in a form that is relevant and appropriate to the audience.

From stage 4 onwards, Noblit and Hare suggest three strategies for relating and synthesising the studies:

1. Reciprocal translation (when accounts are directly comparable);
2. Refutational translation (when the accounts are oppositional); and
3. A line of argument (inference). This last strategy aims to 'discover a "whole" among a set of parts' [7] (p63). This may use Glaser and Strauss' [8] grounded theory approach, or clinical inference: having translated the studies into one another, the line of argument is developed by examining similarities and differences between cases to integrate them in a new interpretation that 'fits' all the studies.

Noblit and Hare make it clear that undertaking meta-ethnography requires previous experience in qualitative methods, if not ethnography per se. They note the influence of the researcher in the process. His/her values and readings of the studies influence the synthesis, which inevitably provides just one interpretation of what are already interpretations of interpretations. The inductive nature of the process means it is emergent, the initial question or area of interest may be adapted or redirected and there are numerous judgement calls along the way[9].

**Examples of meta-ethnography**

To date, this approach to synthesis has only been used for qualitative synthesis, initially in education[7] and more recently in other fields of health-related research. [10,11] Noblit and Hare developed a 'line of argument' synthesis of a series of qualitative case studies on desegregation in US schools. The case studies were seemingly too particular to be effectively summarised, yet Noblit and Hare were able to reconcile seemingly divergent accounts to produce a more general theory about the circumstances in which desegregation is successful by examining the similarities and differences between the studies as if they were undertaking primary qualitative research.

***What are the strengths of meta-ethnography?***

The main value for policy makers of meta-ethnography is the way it may provide a way of explaining seemingly divergent study findings. It could also, in principle, be used alongside a synthesis of quantitative effectiveness data to help answer questions about why interventions

have differing effects in different settings.  This is similar to one way in which realist synthesis can be used[5].

It is revealing to recall that two of the intended audiences for Noblit and Hare's original book were policy researchers and policy makers. Indeed, the genesis of the approach was  the failure to integrate six commissioned ethnographic case studies of school desegregation funded by US National Institute of Education (Noblit and Hare [7] pp19-20).

**What are the limitations of meta-ethnography?**

Noblit and Hare's original focus was on studies they classed as ethnographic; i.e. entailing intensive research involving observation, interviewing and document review and firmly located within the interpretive paradigm.  Subsequent meta-ethnography [10,11] has shown the applicability of the method to a broader range of qualitative studies (i.e. studies informed by different perspectives and using single methods such as interviews). However, it is not clear how meta-ethnography might be used to synthesise qualitative and quantitative data in a single synthesis. It is conceivable that the texts of quantitative reports could be analysed in the same way as the texts of qualitative reports since both types of research are ultimately expressed in text form. The crucial issue is whether the textual explanations in quantitative reports can ever be treated in a way that would allow the identification of metaphors and concepts comparable to those found in qualitative reports. Many ethnographers would object that the purpose and nature of reports of qualitative studies are quite different from quantitative reports, particularly the degree of conceptualisation in the former, and that the enterprise is fruitless.

While the potential of this method for a full integration of qualitative and quantitative studies may be limited, it does offer a way of synthesising existing qualitative research which may be incorporated into the policy decision making process, perhaps alongside a comparable review of the quantitative evidence on a topic.

**2.2    Qualitative cross-case analysis**

*What is it?*

Case studies are used to understand complex social phenomena. Research using a case study approach may be based on a single or multiple cases, and can include a mixture of qualitative and quantitative evidence.  It 'need not always include direct, detailed observations as a source of evidence'[12] (p15), and thus provides a potential method for qualitative-quantitative synthesis. Yin[12] argues that multiple sources of evidence strengthen the case study, allowing the investigator to capture a broader range of issues and allows triangulation (which Yin appears to

understand as convergence on a fact). He argues that research will be more convincing if several sources converge on a fact, as may occur in a synthesis.

Miles and Huberman [13] describe a range of approaches to qualitative data analysis within a single study, which they call 'cross-case analysis/synthesis' (p172). In this context, a case can be a single example, e.g. a patient or an incident, but the approaches they describe have the potential to be used across multiple studies for the purpose of synthesis. In the main, these approaches are ways of presenting findings from a number of cases/studies in new ways to increase generalisability. Miles and Huberman make a useful distinction between case (i.e. a 'whole') and variable-oriented (i.e. some part or aspect of a case) strategies for analysis. Variable-oriented strategies entail looking for themes that cut across cases/studies (thereby inevitably separating the features of the case from its context). Case-oriented strategies include the meta-ethnographic approach described above and the various pattern matching techniques described by Yin [12, 14].

Yin also describes explanation building, a narrative form of pattern matching in which the aim is to develop an explanation using an iterative, constant comparative type approach. By clustering cases, much as in primary research, it is possible to develop types or families and these may further be sorted on key dimensions (e.g. responders to an intervention and non-responders). Yin says this approach is applicable to multiple case studies (e.g. as in Derthick's study of federal housing programs [15]) and thus it has the potential to be used for policy synthesis. Denzin [16] uses a variation of this pattern matching approach by looking for multiple exemplars in an interpretive synthesis, collecting multiple cases that embody a particular concept, thereby identifying essential elements which can then be reassembled into a whole.

A combination of case- and variable-oriented strategies appears to be the most productive as this preserves the context of each case whilst providing a bigger picture of the patterns created by the variables. Miles and Huberman [13] use a meta-matrix approach for qualitative data analysis. They run through a number of ways of interrogating data by displaying it graphically in charts (matrices). Their procedure entails first examining each whole case in detail, then developing case-based matrices to display specific variables from cases, then displaying the cases together in a meta-matrix to allow systematic comparisons. Faced with an unwieldy data set the analyst can standardise and reduce the data by sorting, quantifying and collapsing them into analytic categories which can be displayed together on larger meta-matrices (in essence, theme x case charts). Matrices of varying levels of complexity are described – ordering by cases (looking at

one variable, and ordering all the cases), looking at two variables together and pulling out multiple exemplars for comparison and so on. In all, some 27 different ways of constructing matrices are presented each of which provides different viewpoints on the data.

Two other ways of displaying data are worth considering for qualitative and quantitative synthesis. One is based on generic narrative models that preserve the narrative sequence of cases whilst also looking across their variables which was developed by Huberman[17] as composite sequence analysis. Gladwin's analysis[18] of causal models of decision making is very similar. It uses decision trees and algorithmic charts to show the stages of decision making for each case. Miles and Huberman also describe time-ordered meta-matrices which are a variant of their earlier meta-matrices ordered by time. These charting approaches could allow the analysis of the temporal development of evidence, but it is less clear how they could be used to integrate qualitative and quantitative evidence; i.e. how you might extract data and convert into equivalent text or number formats to draw the chart. Composite sequence analysis appears to be a more complex method of displaying the data but relies on having the data from a number of cases in a standard format, again raising the difficulty of how to transform different types of data to a common equivalent. Another approach is to use a scatter plot, scaling cases on a continuum for each variable and transforming this into a graphic display. Again it is not clear how to transform qualitative and quantitative data to be equivalent, and how one would choose the scale on either axis of the scatter plot.

Moving beyond these methods, which are essentially different ways of displaying data, towards synthesis (in their case theory or explanation building), Miles and Huberman[13] suggest writing notes about conclusions drawn from observation of the matrices, checking out the conclusions against the raw data from the original studies and verifying conclusions by looking for rival explanations.

Ragin's comparative analysis[19] is a further variant of cross-case analysis. It is a qualitative technique in which a matrix (a 'truth table') is developed using Boolean algebra which includes all logically possible combinations of the presence or absence of a set of independent categorical variables and a corresponding outcome (dependent) variable. The data (which in the context of a review would be a set of studies rather than cases within a single dataset) are then searched to find actual cases that match each of the cells in the matrix. Unfilled or logically inconsistent combinations are excluded from the matrix leaving a parsimonious, logically consistent model of the combinations of variables associated with the outcome of interest. This approach is particularly useful for assessing competing explanations for the same phenomenon

and has the advantage, in principle, of allowing the integration of qualitative and quantitative findings since variables are expressed in simple binary form (present-absent). However, it has not been widely used.

**What are the strengths of qualitative cross-case analysis?**

Cross-case analysis offers a potential method for summarising and displaying evidence from multiple qualitative and quantitative studies for combined analyses. Miles and Huberman's descriptions focus on using the method in the context of a primary qualitative study that has multiple cases (e.g. multiple study sites or pilot projects in a single evaluation). However, there does not seem to be anything intrinsic to Miles and Huberman's approach to suggest that it could not be used to display qualitative and quantitative evidence side by side – for example, they imply that qualitative data might be summarised quantitatively in the mega-matrices, and, at the higher levels of complexity, they illustrate how cases can be summarised in binary form on key variables.

Yin[20] provides an example of a primary study combining qualitative and quantitative evidence, in a manner that might be applicable to qualitative-quantitative synthesis. This looked at the research proposal process across 20 universities. The study identified four (qualitative) patterns in the ways universities processed proposals, and quantitative differences in costs associated with the process – looking at these together it was found that high volume proposal producing universities had higher costs (contrary to expectation).

**What are the limitations of qualitative cross-case analysis?**

Yin is clear that analysis of case study data requires experience, and is not for the novice. Miles and Huberman's description of case study methods, whilst supplying more detailed technical instruction addressed to a broad research community including the novice qualitative researcher, also suggests that this approach is complex.

**3. Quantitative approaches**

**3.1 Quantitative case survey**

***What is it?***

The case survey method is a formal process for systematically coding data from a number of qualitative cases that is sufficient for quantitative analysis[21]. Pelz[22] describes such an approach. A set of structured questions is used to extract data from individual case studies which are then treated as observations within a single dataset. Data are then converted to quantitative form for

statistical analysis of the associations between different variables. Thus the case survey is explicitly a way of turning qualitative studies into quantitative data for analysis, thereby allowing an integrated qualitative-quantitative synthesis to be undertaken.

### What are the strengths and limitations of quantitative case surveys?

Of the approaches to case studies described here, a case survey is potentially the most useful for synthesis since it was specifically designed to handle multiple case studies. It allows the aggregation of findings across a series of studies. When there are large numbers of case studies quantitative techniques can easily be used, including meta-analysis. However, case survey is argued to be reductive by some qualitative researchers, despite the fact that characteristics of the context and so on of each case can be extracted and coded for use as explanatory variables.

### 3.2 Content analysis

### What is it?

Content analysis is a systematic technique for categorising data into themes and then counting how often each category or theme occurs in order to identify dominant findings and, thereby, make some generalisations. Though it was developed for primary research on a wide variety of mainly textual information there is no reason to believe that it could not be applied to the synthesis of findings from different studies. According to Stemler[23], content analysis is defined as 'a systematic, replicable technique for compressing many words of text into fewer content categories based on explicit rules of coding.' Content analysis is essentially a quantitative method since all the data are eventually converted into frequencies, though qualitative skills and underlying theory may be needed to identify and characterise the categories into which findings are to be fitted. Content analysis adopts conventional concepts of validity and reliability in that categories have to be defined sufficiently precisely a priori for many assessors to be able to code the same data (e.g. part of an oral transcript of a television interview) in the same way. As a result, categories must be exclusive and exhaustive.

In the context of reviews of qualitative and quantitative research for policy and management, content analysis could be useful in making use of a range of qualitative information in an essentially quantitative analysis. This would have to be undertaken with care since, for example, the number of times a concept is mentioned in the report of a qualitative study partly depends on the extent to which the report comprises quotes from participants (who may not use the term at all) versus the researchers' interpretation and discussion of what participants were saying. Despite this, content analysis could be very helpful in assessing the likely acceptability of a

range of different policy solutions to a problem by determining the degree of media support for each and the most frequently voiced criticisms.

**What are the strengths of content analysis?**

Content analysis has the advantage of being well developed and widely used (e.g. most notably in media studies, especially of political bias), with software available to help in undertaking analyses. As a result, it is possible to handle large amounts of text straightforwardly and systematically in a short space of time. The steps in content analysis can be easily described making it a transparent and replicable technique. The results are also relatively easy and economical to present consisting of tabulations of frequency counts. The results can be fed into a variety of statistical analyses.

**What are the limitations of content analysis?**

However, content analysis can be criticised for being too reductive and for emphasising those phenomena that are amenable to being counted rather than those that are significant interpretively. For instance, while content analysis is good means of determining how often a politician makes reference to a particular thinker, it cannot so easily be used to identify the importance of that thinker in shaping the politician's thinking. However, coding in content analysis can be far more sophisticated than simple counts of individual words.

**4. Bayesian meta-analysis and decision analysis**

*4.1    The Bayesian approach*

*What is it?*

Bayesian thinking offers one of the most attractive means for quantitative analysts to incorporate qualitative research data and other evidence into syntheses and analytical models designed to assist in policy and management decisions.  The basic idea behind Bayesian methods can be described in the following way. If a conventional clinical trial were carried out to find out by how much a new intervention a was superior or inferior to an existing intervention b for the same condition, the statistical analysis would yield, as summary results, a P-value, an estimate of effect and a confidence interval around the estimate. A Bayesian analysis would supplement this by focusing not just on the question, 'what is the effect of intervention a versus intervention b?', but further on the question, 'how should this trial change our opinion about this effect?' This compels the analyst to:

1. State a reasonable (defensible) opinion on the effect of a (the new treatment), excluding evidence from the trail (the 'prior distribution');

2. State the support for different values of the effect of a, based solely on the data from the trial (the 'likelihood'); and

3. Combine these two sources to produce a final opinion about the effect of a (the 'posterior distribution').

The 'posterior' distribution is produced using Bayes' theorem which states that the posterior distribution is proportional to the product of the 'prior' times the 'likelihood'. For those wanting more detail, Spiegelhalter et al. [24] provide an accessible overview of principles and applications of Bayesian methods to health technology assessment.

The Bayesian approach is thus, 'the explicit quantitative use of external evidence in the design, monitoring, analysis, interpretation and reporting of a study'[24]. From this definition, it can be seen that the approach can also be applied to meta-analysis and other forms of quantitative synthesis (see below). The rationale for Bayesian methods in health and health care derives from an awareness of the limitations of traditional trial methods (e.g. the complications for trial design generated by a desire to report multiple sub-group analyses and the attendant risk of type I error (wrongly rejecting a true null hypothesis)), together with the fact that evidence from multiple sources usually needs to be combined to inform a policy decision. For example, in undertaking trials, Bayesians would allow inferences from qualitative studies about how users react to different types of therapy to inform the 'prior' distribution of outcomes as well as using such information to make recommendations about the use of a treatment trialled at a population level on specific types of individuals. It thus allows the use of a range of evidence in addition to formal quantitative trials, including qualitative data, clinical consensus statements, expert views, etc. in drawing conclusions from either individual studies or syntheses of studies. Proponents argue that, as a result, a Bayesian approach is more likely to provide conclusions from research in a suitable form for making clinical or policy decisions[25]. It likely to be particularly useful when data from primary research is weak or lacking and decisions nonetheless need to be taken in an informed and reasonable way.

Bayesian analysts explicitly bring subjective judgement into the conduct and interpretation of scientific research, and evidence-based decision making on the grounds that ultimately decisions depend on such judgements. They then attempt to identify what it is reasonable for an observer to believe in light of the available data. By putting the subjective element into the open, it is argued that it is more likely to be amenable to rational discussion and control. This also means that a Bayesian analysis explicitly takes account of the perspective of the potential user of the analysis (e.g. whether these are pharmaceutical companies, regulators, payers,

physicians, patient representatives, etc.) and draws attention to the fact that the implications of scientific research for making decisions depend on the perspective(s) adopted. Such an approach has much in common with widely accepted views among qualitative researchers that different pictures of reality will be produced by collecting data from different groups and using different methods.

A Bayesian approach can be used in a number of different forms of synthesis where different forms of evidence need to be brought together, including cross-design synthesis (also known as 'grouped meta-analysis'), Bayesian meta-analysis, and comprehensive decision modelling or decision analysis. Cost-effectiveness analysis also lends itself to a Bayesian approach. Health economists are increasingly turning to Bayesian methods, particularly when cost-effectiveness analysis is to be used directly as an input to a specific decision, and also for sensitivity analysis. These different 'applications' of a Bayesian approach are briefly described below.

### 4.2 Cross-design synthesis

This is a form of meta-analysis, which allows the mixing of different quantitative research designs (e.g. RCTs and observational studies) and the pooling of evidence using modelling to estimate a 'true' effect of a policy or programme, conditional on both the design of the study and on the characteristics of the relevant population[26 27]. The method is based on making explicit statistical adjustments to studies and modelling their likely biases in order to produce a 'true' overall estimate of effect. Instead of eliminating low quality studies from the synthesis, such studies are used where they can provide information that compensates for weaknesses or gaps in the high quality research available. For instance, many RCTs are of restricted applicability because they tend to be carried out on unrepresentative populations, so wider-ranging database studies can be used to supplement the synthesis dataset as long as their potential biases are explicitly allowed for in the modelling of effects. Hierarchical modelling is used to allow for quantitative within-and-between-sources heterogeneity in the pooled dataset.

Although not a Bayesian approach per se, cross-design synthesis lends itself to Bayesian methods. It allows for a priori beliefs regarding qualitative differences between various sources of evidence to be included in the analysis. These subjective judgements are likely to be informed by a wide range of qualitative and non-research information and are likely to be context-specific. As a result, the goal of cross-design synthesis is not to produce a universally applicable answer, but one that is 'true' in the circumstances.

Though cross-design synthesis has been applied to questions of effectiveness, this integrated form of analysis using different databases, case-control and prospective studies, leads naturally into both cost-effectiveness analysis and comprehensive decision modelling with the addition of evidence on 'utilities' and costs (see below).

## 4.3    Bayesian meta-analysis

A Bayesian approach can be applied to the more familiar meta-analysis of RCTs and other quantitative studies, thereby allowing both qualitative and quantitative data to be used together (Box 3 in the main paper contains a detailed example of the approach).[28]  Qualitative studies are used to inform the prior probability distribution and help identify relevant variables to include and their likely effects. The prior distribution is then combined with the quantitative studies to produce the overall synthesis. The qualitative studies contribute in a way similar to their use in much conventional health research by identifying the variables of interest for the subsequent quantitative meta-analysis (i.e. the approach is perhaps more sequential than transformative of the original data), including possibly some variables or factors not included in any of the quantitative studies. The prior probabilities of a variable being important are modified in the analysis of the relevant quantitative studies to produce an overall, posterior distribution of probabilities.  This approach to meta-analysis has the disadvantage of being difficult to accomplish well and may be criticised by qualitative researchers for only allowing the qualitative evidence to inform the prior assumptions about variables and their effects.

## 4.4 Bayesian approaches to cost-effectiveness analysis

Increasingly, health economists argue that Bayesian methods are crucial for useful cost-effectiveness analysis[29] on the grounds that hypothesis testing is of limited relevance in economic evaluations since additional, non-trial evidence is needed to produce an estimate of the cost-effectiveness of a policy or intervention to inform a decision. This has become particularly apparent in decisions on the regulation of pharmaceuticals and medical devices where cost-effectiveness estimates of the potential advantages of innovations over existing technologies and drugs frequently have to be made ahead of definitive research findings and/or when the innovation itself and its application are still being refined (see Box 1 for an example). For similar reasons, a Bayesian approach is an attractive way of providing a quantitative estimate of the likely cost-effectiveness of undertaking research studies. Such 'value of information' studies need to take into account the likely impact on clinical practice of the results of trials and in order to do so, need to be able to estimate the probability of different trial results occurring.

**Box 1: An example of a systematic review and economic decision modelling for the prevention and treatment of influenza A and B[28]**

*Objectives*
- To establish the clinical and cost-effectiveness of available drugs for the treatment of influenza relative to the existing method (no treatment or antibiotics)
- To establish whether two of the available drugs are effective and cost-effective alternatives to the existing method of prevention (no intervention or vaccine)
- To make policy recommendations

Methods
Systematic review and meta-analysis of randomised trials to look at effectiveness including additional evidence from pharmaceutical companies not available from the published literature. Separate reviews of evidence of effectiveness in population sub-groups (children, healthy adults, 'high risk' (over 65 years with concomitant disease) and elderly people in residential care)

Economic decision models constructed to examine cost-effectiveness and cost-utility (marginal cost per QALY gained) of a range of feasible strategies for treating and preventing influenza, informed by the systematic reviews and other information (e.g. on likely timing of treatment after onset of symptoms in real world settings)

Estimates of the probability that costs per QALY lie within particular ranges for each strategy and for population sub-groups

Sensitivity analysis of results

*Features*
- Analysis focused on comparing a range of feasible potential treatment and prevention strategies
- Not all comparisons had been directly studied before, but were relevant to policy decisions
- Analysis using UK-based or adjusted overseas information to produce UK-relevant recommendations (i.e. context-specific)
- Comparisons focused on marginal costs and benefits (i.e. assuming current policies exist)
- Range of research (RCTs and database studies) and non-research (expert opinion) data brought together, spanning many decades
- Results expressed in terms of their probabilities

Conclusions
Cost-effectiveness varies between intervention strategies and target population sub-groups. In all cases, the cost-effectiveness ratios for vaccination were either low or cost-saving. Cost-effectiveness ratios of antiviral drugs were relatively unfavourable except for some scenarios involving treatment of elderly people in residential care where antivirals as an additional strategy could be cost-effective (i.e. 60% likelihood of cost per QALY gained below 30,000 pounds). There were a number of areas where further research could improve the modelling of cost-effectiveness.

---

A Bayesian approach can also be taken to conventional sensitivity analysis in which prior probability distributions are placed over the uncertain inputs to the analysis and the resulting distribution of potential cost-effectiveness ratios is generated by simulation.

## 4.5 Comprehensive decision modelling

A Bayes informed cost-effectiveness analysis can be developed into a full-scale decision-theoretic model[31]. This is a formal, analytical means of incorporating a wider range of evidence

beyond research studies into a synthesis as well as making explicit the value judgements necessary to identify the best course of action for decision makers (with the aim of increasing the likelihood that decisions will be informed by evidence).

Comprehensive decision modelling incorporates all the major steps in a rational decision process (see Box 2) – synthesis of the available scientific evidence from all levels of the evidential hierarchy, valuation of the outcomes in terms of 'utilities' (e.g. quality-adjusted life years gained) and preference elicitation (defining and measuring the trade-offs between policy goals such as maximising health gain versus maximising access improvements among low users). The analysis uses Bayesian statistics together with stochastic cost-effectiveness analysis and is typically implemented using Monte Carlo simulations. It attempts to compare the costs and benefits of different programmes or interventions allowing for the uncertainty underlying the available evidence so that results can be presented to decision makers as probabilities that each course of action or intervention is the most cost-effective. The modelling can incorporate the impact on outcomes and costs of a mix of policies or programmes operating at different levels in a system such as the simultaneous impact on smoking rates of changes in excise duty (macro), smoke-free workplace legislation (meso) and health promotion initiatives (micro).

**Box 2: Stages in comprehensive decision modelling**

| | |
|---|---|
| Synthesis of best available research evidence | Information on effects and costs of different policies/programmes relevant to the particular context/population and 'weighted' according to their probability of being 'true' |
| Valuation studies | Valuing outcomes (e.g. using time trade-off, standard gamble, etc.) to produce 'utilities' |
| Preference elicitation | Define and measure trade-offs between goals and outcomes from views of population (ideally from specially collected data from surveys and focus groups, but can use any other intelligence available) |
| Link research evidence to population preferences and valuations | |
| Assess costs and benefits of options and state probability of each occurring in the way predicted | |

The quality of the modelling rests on careful searching and processing of evidence relating to each parameter in the model, including explicit valuation of the outcomes (if there are more than one) and costing of each policy option. Each parameter is usually presented as a probability distribution to represent the inherent uncertainty underlying the evidence used.

Proponents argue that by making explicit the value judgements and trade-offs inherent in decisions, as well as the effect of variations in evidence quality and gaps in evidence, the approach is superior to traditional, non-analytical, implicit decision making processes which risk

overlooking important issues altogether without it being possible to know. Dowie coins the phrase 'Taking Into Account and Bearing in Mind' to summarise the conventional, largely implicit approach to decision making and contrasts it with Bayesian decision analysis.[31] He argues that comprehensive decision modelling may still be imperfect, but is likely to be superior to the decisions of committees (e.g. Cabinet sub-committees) since, paradoxically, it deals with all the uncertainties and trade-offs explicitly, transparently and quantitatively rather than implicitly, covertly and qualitatively.

**What are the implications of Bayesian approaches?**
The use of Bayesian decision modelling has a number of major implications for the way in which policy makers relate to those undertaking syntheses for policy and management purposes. Firstly, it makes little sense for policy makers to commission the creation of a comprehensive decision analytic model unless they are prepared to use it as a direct guide to their eventual decisions. If such a model incorporates all available information and includes values and preferences that are broadly regarded as reasonable (particularly by the decision makers), then, the proponents argue, the output should reflect the best course of action in the circumstances from the perspective of those taking the decision. Accordingly, it makes no sense to use the output from the modelling as yet another input to a traditional implicit, 'taking into account and bearing in mind' decision. This is a radical shift in decision-making processes that many policy makers are reluctant to embark upon.

The second major implication of attempting to bring together all relevant scientific and other information is that policy makers have to define and explain up-front the key value judgements and trade-offs that they are prepared to make. For example, the relative weight to be given to different outcomes (e.g. reducing child poverty versus encouraging people into paid work in a welfare reform) has to be determined so it can be built into the decision analytic model. Population preferences may well be in conflict or even incoherent, in which case, judgements will have to be made by policy makers that the modellers can work with.

Thirdly, and relatedly, policy makers have to make available to the analysts (who may be in-house staff or external) all the existing information that is relevant as well as funding the collection of extra data such as population preferences if time and resources permit. In systems where the policy process has traditionally been confidential, the requirement for sharing information with the analysts represents a major change.

The fourth implication relates to the analysts. There will no longer be any neat cut-offs between evidence of acceptable scientific quality and that which can be excluded. In addition, it is no longer sufficient for the analyst to conclude that the evidence is inadequate and leave it to someone else to cope with the decision making consequences. The analyst/synthesiser has to produce a best solution in the circumstances with the information available.

Fifthly, qualitative views, preferences and judgements have to be quantified whether from research or 'intelligence' collected more informally (e.g. from stakeholder management processes). Qualitative phrases such as 'giving x due weight', 'establishing the right balance between x and y', 'group a is strongly supportive of objective x' 'x is highly likely to be the reaction of group b to policy n', and so on, all carry quantitative and/or probabilistic implications which have to be drawn out explicitly for modelling to be possible. This is not an easy task.

Finally, a Bayesian would argue from the outset that the conclusions of all decision analyses depend on who is conducting them, for what purpose and on the basis of what evidence and opinion (i.e. context is vital). As a result, it is not possible, or desirable, to portray the eventual decision as the only 'true' one, only that this is the best decision given all the material to hand to construct the model. For some policy makers, this is an uncomfortable position to be in since it means explicitly presenting the decision in terms of uncertainty rather than the usual assumption of certainty. This has large implications for the conduct of politics as well as government processes.

**What are the strengths of Bayesian meta-analysis and decision analysis?**

Perhaps the strongest aspect of the above methods is that they allow the synthesis of all available sources of evidence from RCTs, databases, professional consensus exercises, expert opinion, tacit knowledge, population focus groups, other qualitative research, etc. into a single quantitative model. Beliefs about the differences between these different sources of evidence are explicitly included in the research synthesis. They try to make explicit and transparent all the judgements that have to be made to assess options and take decisions in specific contexts. They attempt to quantify the effects and costs of each potential option on the same basis. This should, in principle, improve comparison across options within a single decision, but also improve consistency across separate decisions (e.g. this may improve the allocative efficiency of government).

Models can be updated at any time as new evidence and information on any part of the model becomes available, thus permitting decisions to be revised after a period of time. Models also

show clearly where the evidence to support decisions needs to be improved and can begin to show the value of new research.

Bayesian decision analyses admit to the inherent error and uncertainty of decisions since they relate, by definition, to unknowable future states. This approach strengthens accountability for decisions since it is possible for observers to assess whether the decision taken was reasonable given the parameters of the model and can propose improvements which can be discussed. It militates against 'off the cuff', poorly informed decisions.  It allows explicit recognition of multiple perspectives, the perspective(s) from which the analysis was undertaken and how the results would alter if other perspectives were included. This resonates with the view of qualitative researchers who tend to argue that who you ask and where you sit influence what you find and the conclusions you draw.

**What are the limitations of Bayesian meta-analysis and decision analysis?**

Decision analysis and Bayesian approaches raise concerns about the fact that studies with 'weaker' designs will be included in the synthesis, thereby potentially undermining the validity of the analysis from certain perspectives. There are concerns about feasibility; for example, that the biases and differences in rigour between studies are, in practice, tricky to handle in order to produce a synthesis of effects, or that even the best model cannot cope with the presence of multiple stakeholders with different utility functions which are hard to tease out and to reconcile. The risk is that the model gives an impression that we know more than we really do. These models could produce a perception that the subtleties of policy advice and decision-making are being reduced to mechanistic formulae even if this is not necessarily the case.  They can be accused of undermining the role of ordinary policy makers by putting decisions in the hands of unaccountable 'experts' in modelling. This could be seen as undemocratic, particularly if the results conflict with politicians' 'priors'.  Relatedly, the complexity or specialist nature of the processes for developing the models and the models themselves can be difficult to communicate to lay audiences.  Models require a considerable amount of work to assemble (e.g. to convert qualitative information into utilities and probabilities) and there may not be time or analytical capacity to do this.  Finally, decision analytic approaches to using research synthesis for policy and management decision making may require unrealistic, revolutionary change in the way that public policy is made to be useful.

The advocate's response to most or all of the above criticisms and concerns is that they implicitly compare decision analytic approaches to a perfect approach that does not exist. For the

advocate of these approaches, the argument can only be resolved through using decision analytic models and assessing the consequences of the decisions they indicate.

**References**

1. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A.  Synthesising qualitative and quantitative evidence: a review of possible methods. *Journal Health Services Research & Policy* 2005; 10: 45-53

2. Pawson R, Bellaby P.  Realist synthesis: an explanatory focus for systematic review. In: Popay J, ed.  *Putting effectiveness into context: methodological issues in the synthesis of evidence from diverse study designs.* London: Health Development Agency, forthcoming.

3. Rutter, M., Maughan, B.,Mortimore, P., Ouston, J. *Anti-social Behaviour by Young People*.  Cambridge: Cambridge University Press, 1998

4. Shadish W., Cook T, Leviton L. *Foundations of program evaluation*. Newbury Park: Sage, 1991

5. Pawson R, Greenhalgh T, Harvey G, Walshe K. Using realist methods to produce syntheses of evidence for use by managers and policy makers.  Ottowa/London: Canadian Health Services Research Foundation/NHS Service Delivery and Organisation Research Programme, 2005, forthcoming

6. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Roen K, Rodgers M.  *Guidance on the conduct of narrative synthesis in systematic reviews.*  Draft report from ESRC Methods Programme.  Lancaster: Institute for Health Research, University of Lancaster, 2004

7. Noblit G, Hare R. *Meta-ethnography: synthesising qualitative studies*. Newbury Park CA: Sage, 1988

8. Glaser B, Strauss A. *The discovery of grounded theory: strategies for qualitative research*. New York: Aldine, 1967

9. Light RJ, Pillemer DB.  *Summing up: the science of reviewing research*.  Cambridge, Mass: Harvard University Press, 1984

10. Campbell R, Pound P, Pope C, Britten N, Pill R, Morgan M, Donovan J. Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. *Soc Sci Med* 2003;56:671-684

11. Britten N, Campbell R, Pope C, Donovan J, Morgan M, Pill R. Using meta-ethnography to synthesise qualitiatve research: a worked example. *Journal of Health Services Research and Policy* 2002; 7: 209-15

12. Yin R. *Case study research, design and methods.* Applied Social Research Methods Series Vol 5 Thousand Oaks CA: Sage, third Edition, 2003

13. Miles MB, Huberman AM. *Qualitative data analysis: an expanded sourcebook.* London: Sage, 1994

14. Yin R. *Case study research, design and methods.* Applied Social Research Methods Series Vol 5 Thousand Oaks, Sage, 1984.

15. Derthick M. *New towns in-town: why a federal program failed.* Washington DC: The Urban Institute, 1972

16. Denzin N. *Interpretive interactionism.* Applied Research Methods Series 16. Newbury Park CA: Sage, 1989

17. Huberman A. The professional life cycle of teachers. *Teachers College Record* 1989;91:31-57

18. Gladwin C. *Ethnographic decision tree modelling.* Qualitative Research Methods Series No.19 Newbury Park, CA: Sage, 1989

19. Ragin C. *The comparative method: moving beyond qualitative and quantitative strategies.* Berkeley: University of California Press, 1987

20. Yin R. *Applications of case study research.* Applied Social Research Methods Series Vol 34. Thousand Oaks, Sage, second edition, 2003

21. Yin R, Heald K. Using the case survey method to analyse policy studies. *Admin Sci Q* 1975;20:371-81

22. Pelz D. Use of innovation in innovating processes by local governments. Ann Arbor: University of Michigan, CRUSK Institute for Social Research, 1981

23. Stemler S. An overview of content analysis. *Practical Assessment, Research and Evaluation* 2001; 7(17). Retrieved March 11 2004 from http://PAREonline.net/getvn.asp?v=7&n=17.

24. Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Bayesian methods in health technology assessment: a review. Health Technology Assessment 2000; 4(38)

25. Dixon-Woods M, Fitzpatrick R, Roberts K. Including qualitative research in systematic reviews: problems and opportunities. *Journal of Evaluation in Clinical Practice* 2001: 7: 125-133

26. Droitcour J, Silberman G, Chelimsky E. Cross-design synthesis: a new form of meta-analysis for combining results from randomised clinical trials and medical-practice databases. *International Journal of Technology Assessment in Health Care* 1993; 9: 440-9

27. US General Accounting Office. *Cross-design synthesis: a new strategy for medical effectiveness research.* Washington DC: General Accounting Office, 1992

28. Roberts KA, Dixon-Woods M, Fitzpatrick R, Abrams KR, Jones DR. Factors affecting the uptake of childhood immunisation: a Bayesian synthesis of qualitative and quantitative evidence. *Lancet* 2002: 360: 1596-1599

29. Luce BR, Claxton K. Redefining the analytical approach to pharmacoeconomics. *Health Economics* 1999; 8: 187-9

30. Turner D, Wailoo A, Nicholson K, Cooper N, Sutton A, Abrams K.  Systematic review and economic decision modelling for the prevention and treatment of influenza A and B. *Health Technology Assessment* 2003; 7(35)

31. Dowie J. Towards value-based, science-informed public health policy: conceptual framework and practical guidelines. WHO/Department of Health Consultation on Risks to Health: Better Management for Decision Making, 2001