A realist evaluation of a normal birth programme

Purva Abhyankar *[1] PhD, MSc.* ***Helen Cheyne*** *[2] PhD MSc, RM, RGN.* ***Margaret Maxwell*** *[3] PhD, MA.* ***Fiona Harris*** *[4] PhD, MA.* ***Christine McCourt*** *[5] PhD, BA.*

1. Research fellow, Nursing Midwifery and Allied Health Professions Research Unit, University of Stirling, Stirling FK9 4LA Scotland. Email: purva.abhyankar@stir.ac.uk
2. Professor of maternal and child health, Nursing Midwifery and Allied Health Professions Research Unit, University of Stirling, Stirling FK9 4LA Scotland. Email: h.l.cheyne@stir.ac.uk
3. Professor of mental health and deputy director, Nursing Midwifery and Allied Health Professions Research Unit, University of Stirling, Stirling FK9 4LA Scotland. Email: margaret.maxwell@stir.ac.uk
4. Research fellow, Nursing Midwifery and Allied Health Professions Research Unit, University of Stirling, Stirling FK9 4LA Scotland. Email: fiona.harris1@stir.ac.uk
5. Professor of maternal and child health, School of Health Sciences, The City University London, Bartholomew Close, London EC1 7PN England. Email: christine.mccourt.1@city.ac.uk

This paper is based on a keynote lecture given by Professor Helen Cheyne at the first meeting of the Scottish Branch of the Doctoral Midwifery Research Society on 17 May 2013.

Abstract

Background. Complex programmes are characterised by multiple components, acting independently and interdependently, at multiple levels, within diverse and dynamic systems. High-quality evaluation of such programmes is imperative for optimising their development, implementation and effectiveness, but is often challenging. There is debate about whether the traditional outcome-focused approaches are feasible or appropriate for evaluations of large-scale, complex programmes as they often fall short of explaining how and why they do or do not work. Theory-driven approaches offer a more appropriate alternative as they attempt to uncover the black-box between the programme's inputs and the resulting complex pattern of outcomes.

Aim. This paper illustrates drawing upon a realist evaluation to assess a complex programme to support normal birth.

Method. Firstly, the programme theories – the ideas about how the programme would bring about change – were elicited from programme developers and key stakeholders. Secondly, these initial hypotheses were tested out by collecting data on how the programme worked in different contexts, using a multiple case study design. Thirdly, the data were analysed and interpreted to refine the programme theories in light of evidence on how the programme unfolded in practice.

Findings. Described in detail are the process of conducting a realist evaluation, methods used, steps in data analyses, challenges encountered and the approach adopted to overcome them. The usefulness of this approach and some limitations are discussed.

Key words: Realist evaluation, normal birth, complex interventions, programme evaluation, evidence-based midwifery

Introduction

Those involved in the planning and delivery of health care are increasingly aware of the importance of demonstrating that service innovation and change 'works'. However, in the context of the fast pace of change in a real world, health service researchers may be challenged to conduct rigorous research and provide timely and robust information to support the delivery of high-quality health care. If the concepts of rigorous and timely are not to be incompatible, a number of factors need to be considered at the early stage of study design.

Traditionally, the focus of evaluative inquiry has been dominated by summative or outcome-focused approaches, which aim to assess the impact and effectiveness of interventions in achieving intended outcomes (Greenhalgh et al, 2004; Pawson and Tilley, 1997). Within this tradition, randomised controlled trials and other experimental designs have been regarded as the gold standard for assessing the worth of single or multi-component healthcare interventions (MRC, 2000). This approach is also advocated in evaluations of complex programmes, which are often characterised by multi-level components acting independently as well as interdependently (Craig et al, 2008; Campbell et al, 2000). However, whether this approach is feasible or appropriate for evaluations of complex programmes is highly contested, with some emphasising the assessment of 'impact' and others focusing on the 'process'. In recent years, it has been increasingly acknowledged that, in many situations, it is not possible to directly answer or appropriate to ask the question, 'does this programme work?' (Mackenzie et al, 2010). Two considerations are important in deciding the right questions to ask. One is concerned with the nature of complex programmes and the other the contexts in which they are implemented.

Firstly, complex healthcare programmes often comprise multiple components operating at multiple levels into systems that are themselves complex, dynamic and diverse. Programmes may comprise one or more interventions that have been tested for efficacy in a trial and now require to be implemented in real world settings. Several such components/complex interventions may be combined and implemented across diverse contexts in anticipation that these combined efforts will result in a greater impact than those achieved by individual interventions. In many cases, programmes may already be in place before evaluation is considered, or they may be required to be rolled out across all areas simultaneously, giving little opportunity for a control group. In addition, the programme components are likely to interact with each other in unexpected ways to produce a range of unanticipated and unintended outcomes. Evaluations in such circumstances need to focus not on whether programmes work, but on how well they are implemented, how different components interact, and what makes them work or not. Secondly, complex programmes are implemented into pre-existing social systems that are themselves complex, dynamic and ever-changing. Health services are delivered in organisations that are subject to frequent change at every level of service delivery and are affected by factors such as service reorganisation, changes in health policy direction and clinical guidance, as well as financial constraints, changing demographics, health needs and demands. Interventions interact with local contexts in complex, dynamic and unpredictable ways, leading to a complex and often mixed pattern of outcomes. In such circumstances, there is often little evidence to identify linear, causal relationships between the programme inputs and observed changes in practice or clinical outcomes, as is required by outcome-focused approaches (Greenhalgh, 2009). In evaluating complex programmes, it is more pertinent to ask how the components interact with the context, why they work well in some contexts more than others and how contextual factors relate to impact and implementation (Pawson and Tilley, 1997). Outcome-focused approaches are inherently inadequate for this task as they are not geared up to explaining the 'black box' of how and why programmes work, and fail to acknowledge the active role played by the social systems into which programmes are implemented (Ling, 2012).

Evaluation research, drawn from applied social sciences, offers systematic procedures for undertaking rigorous systematic evaluation to address questions that cannot be addressed by the outcome-focused approaches (Rossi et al, 2004). Two alternatives to the traditional approach are formative and process evaluation (Robson, 2011). Process evaluation is concerned with understanding what actually occurs in the programme and explaining how and why the outcomes are brought about or not. Formative evaluation intends to inform and improve the quality of the development of a programme through a constant feedback of process and outcomes to developers and implementers (Scriven, 1991). In the context of health service evaluation, a similar framework by Donabedian (1980) argues that information about quality of care can be drawn from the 'structure', 'process' and 'outcomes' of care.

The science of evaluation has developed rapidly and several different approaches are now commonly used in healthcare evaluation, including logic models, theory of change and realist evaluation (Blamey and Mackenzie, 2007). Although there are slight differences among these, they all offer a theory-driven framework for evaluation, based on a model or theory about how the programme works and seeks to get inside the 'black-box' of a programme to explain the complex and mixed pattern of outcomes (Pawson and Tilley, 2004; Weiss, 1997).

Realist evaluation

Realist evaluation adopts a distinctive view on the nature of programmes, how they work and what is involved in understanding and explaining them. According to the realists, programmes are social systems where there is a constant interplay between human agency (people's capacity to act freely and shape their lives) and social structures (the environment or circumstances they work in) such that any change is a result of an interaction between individuals and the systems they work in. The social structures provide individuals with resources that enable them to act, as well as placing limits on their behaviour. However, the behaviour of human agents is not exclusively governed by social structures; individuals are also able to transform social structures by responding creatively to the circumstances they find themselves in (Pawson and Tilley, 2004; McEvoy and Richards, 2003). Programmes therefore represent ideas or theories, often stemming from the minds of those who develop and implement them, about the way in which wrongs may be put to right. The task of realist evaluation is to identify the core theories about how the programme is supposed to work and test them out to see if they are plausible, practical and valid.

Realist evaluation contends that it is not programmes that work, it is the people involved in the programme who make them work. It suggests that programmes introduce resources and opportunities for change, but the actual mechanisms that bring about change are located in the reasoning and capacity of those touched by the intervention and contingent on the social context in which they work. Realist evaluation seeks to explain the complex relationship between the mechanisms activated by the programme, the context that influences their workings and the outcomes they produce. Mechanisms refer to the process of how people interpret and act upon the resources offered by the programme and their capacity to put their choices into practice. Context refers to those features of the conditions in which programmes are introduced that enable or disable the operation of the programme mechanisms. Finally, outcomes refer to the intended and unintended consequences of interventions, resulting from the activation of different mechanisms in different contexts. The explanatory proposition of realist evaluation is that interventions work (have successful outcomes – O) only in so far as they introduce appropriate ideas and opportunities (mechanisms – M) to groups in the appropriate social and cultural conditions (contexts – C) (Pawson, 2013; Pawson, 2002; Pawson and Tilley, 1997).

There are three broad phases to realist evaluation (Pawson and Tilley, 2004). The first phase seeks to identify and formalise the programme theory by gathering data from developers and key stakeholders of the programme on how the programme is expected to work. These data are used to build hypotheses about how the programme is expected to work, for who, in what circumstances and to produce what outcomes. These hypotheses are in the form of context (C), mechanism (M), and outcome (O) configurations. In the second phase, data are gathered on the contexts, mechanisms and outcomes in sites where the programme is implemented to interrogate each of these hypotheses. In the third and final phase, the set of context-mechanism-outcome hypotheses are put to test, using the data gathered at phase two. These analyses are aimed at testing if the proposed theory (CMO configurations) explains the complex pattern of observed outcomes. The original programme theories are then refined in the light of data generated in the testing phase, which gives way to middle-range theories, indicating how programmes activate mechanisms among whom and in what conditions to bring about different outcomes.

This paper describes how the realist evaluation approach was used in the evaluation of a Scotland-wide programme of change in maternity care: the *Keeping childbirth natural and dynamic* programme. The evaluation is reported in full elsewhere (Cheyne et al, 2013). This is a discussion of the methods used, the opportunities and challenges encountered.

Abhyankar P, Cheyne H, Maxwell M, Harris F, McCourt C. (2013) A realist evaluation of a normal birth programme. *Evidence Based Midwifery* **11(4)**: 112-119

© 2013 The Royal College of Midwives. *Evidence Based Midwifery* **11(4)**: 112-119

Keeping childbirth natural and dynamic – a national programme

The concept of normal birth as a 'good thing' has gained widespread acceptance in the UK and many developed countries. A growing body of research has indicated practices likely to be effective in supporting normal birth and these practices have been endorsed in government health policies and healthcare guidelines. Despite this, most women experience some form of medical intervention in labour and the caesarean section rate continues to rise. In 2007, the Scottish government introduced *Keeping childbirth natural and dynamic* (KCND), a programme of change in maternity care, with the aim to maximise opportunities for women to have as natural a birth experience as possible, reduce unnecessary interventions in low-risk pregnancy and childbirth and to provide women-centred care.

It was a three-year programme funded and led by the chief nursing officer's directorate (within the Scottish Government Health Directorates). A multidisciplinary national steering group was established to oversee programme development and monitor progress toward targets. The group was chaired by the chief nurse for Scotland and comprised representatives of the main stakeholder groups relevant to maternity care, including the health board executive nurse directors, midwifery service leads, consumer organisations and professional bodies (the Royal Colleges of Midwives, Obstetricians, GPs, Paediatricians and Anaesthetists). The agencies involved in quality and monitoring of health care and education in Scotland were also represented (Health Scotland, NHS Education Scotland, Quality Improvement Scotland). A consultant midwife was seconded to be the national programme director and, in each health board, a senior manager was identified as programme lead with responsibility for reporting back to the national steering group. Central funding was provided for the appointment of a consultant midwife in each health board to facilitate and support the programme. The main objectives were to:

- Ensure evidence-based care, reduce unnecessary intervention, and support informed choice
- Introduce multiprofessional care pathways based on a traffic light risk assessment system; green: midwife-led care, amber: further assessment required, red: maternity team care with flow between levels of care pathways as risk altered
- Initiate the lead maternity care professional designated by risk; the midwife would be the lead for healthy women, those with more complex needs would have maternity team care, led by an obstetrician while the woman's GP had ongoing responsibility for medical care
- Establish the midwife as the first point of professional contact for women in pregnancy. The midwife would undertake early risk assessment and streaming of women to the appropriate care pathway.

Keeping childbirth natural and dynamic – the evaluation

The Nursing, Midwifery and Allied Health Professions Research Unit was commissioned to undertake the evaluation of the KCND programme from 2008 to 2011. This provided the opportunity to observe the unfolding programme as it was developed and implemented into maternity care practice.

The aims and design of the evaluation were shaped by the following considerations. Firstly, KCND was implemented on a national (Scotland-wide) basis and the evaluation took place alongside the implementation. The 'real-time' nature of the evaluation meant that the researchers participated as active observers, influencing the unfolding development of the programme as well as evaluating it. In this situation it is not possible to pre-identify the programme's active ingredients, or ring-fence intervention and control groups in ways necessary to conduct a randomised controlled trial even of complex interventions. This meant that an experimental design was not feasible. Secondly, KCND was a complex healthcare programme having multiple objectives and numerous work streams operating across different levels of the maternity care system, it brought together several interventions or practices for which there was existing trial evidence of effectiveness, but implementation in practice was variable. A method was required that would be sensitive to the contextual factors and provide explanations of why maternity services in some areas had more fully adopted pro-normal practices, while others appeared to be resistant to change from longstanding medical models of care. A realist evaluation was chosen due to its acknowledgement of the importance of context in understanding the way in which programmes work in real-life situations.

The valuation therefore aimed to understand how the KCND programme was implemented, how the different components worked, and how the local contexts shaped the programme's impact on clinical practice and outcomes. A secondary aim was to draw transferable lessons for future national implementation of healthcare strategies.

Realist evaluation provided an overall framework for the evaluation and informed decisions about design and method (Pawson and Tilley, 1997). Accordingly, there were three stages to the project, as described above.

Stage one – identifying programme theory

The objective at the first stage was to understand the purpose of the programme, what it was anticipated to achieve and how this informed decisions about its development and implementation. Semi-structured interviews and focus groups were used to gather data from those involved in developing the programme and other key stakeholders; these data were used to develop the programme theory.

All members of the national steering group were invited to participate in individual interviews. The 14 consultant midwives employed as part of the KCND programme were invited to attend one of two focus groups. Data were collected using a topic guide focusing on the contexts, mechanisms and potential outcomes of KCND. Interviews and focus groups explored the stakeholders' account of the purpose of the KCND programme, the key aspects of the programme, how it would be implemented, how it was expected to work, likely facilitators and barriers to the programme and likely impact on practice. Consultant midwives were additionally asked to discuss their experience of participating in KCND and the strategies they employed to implement and support the programme. Ultimately, 12 stakeholders took part in individual interviews and a total of 13 midwives took part in two focus groups. The interviews and focus groups were audio recorded, transcribed verbatim and managed using the software package QSR NVivo 8.

Analysis

The framework approach (Ritchie and Spencer, 1994) was adopted to analyse the data generated from interviews and focus groups. An initial coding frame was developed using data from three transcripts and the key concepts of the realist evaluation framework – contexts, mechanisms and outcomes. These transcripts were selected to ensure that the main stakeholder groups – such as midwifery, obstetrics and management – were represented. The transcripts were read and re-read independently by two researchers in order to become thoroughly familiar with the data. Data from each of the transcripts were divided into 'units of analysis' – data segments containing discrete bits of information to be assigned to categories in subsequent analytical steps. Each unit was then assigned a meaning, known as a code. At this initial stage, the codes were largely descriptive and merely reflected what the unit suggested with respect to the main topics covered in the interviews, such as programme aims, activities, process, change mechanisms, barriers and facilitators. The codes, first assigned independently by two researchers, subsequently underwent several iterations and discussions until they were refined to accurately describe the meaning contained.

Codes from the three transcripts were then considered together to look for similar codes, which were sorted and grouped under higher order categories. At this stage, the codes were also classified according to the realist concepts of contexts, mechanisms and outcomes. All codes that suggested a change in people's minds and actions (reasoning, feeling, behaviour, attitudes, and beliefs at individual, interpersonal, social and organisational levels) in response to the changes introduced by the programme were classed as mechanisms. For instance, codes suggesting that the programme would work by providing 'local champions of normality', 'maximise midwives' role and accountability' or 'increase midwives' confidence', were grouped together as programme mechanisms. Codes describing the pre-existing enabling or disabling conditions, as well as measures introduced by the programme to support implementation, were categorised as contexts. Codes that described the intended and unintended consequences of the programme, whether behavioural, attitudinal or clinical, were classified as outcomes.

After several iterations and revisions, a preliminary coding framework was developed which was then systematically applied to the remaining transcripts. New categories were added as they emerged from subsequent data. Once all the data had been coded, the coding framework was revisited to search for similarities and differences among categories and identifying higher order themes. The resulting refined framework was used to generate initial hypotheses about what mechanisms would be triggered by the programme, in what contexts to achieve what outcomes. These formed the initial CMO configurations. Below is a discussion of some of the challenges encountered during the process of generating initial CMO hypotheses and the approach adopted to overcome them.

Identifying programme components

The first challenge was in identifying the resources and opportunities offered by the programme to make change happen, which is a crucial first step in identifying the mechanisms triggered by them, for example: identifying how they are interpreted and acted upon by those who are involved in or affected by the programme. In interventions involving single or multiple (albeit inter-related) components that lead to a coherent set of predictable mechanisms and outcomes in fairly well known contexts, this is relatively straightforward. However, with a large and complex programme like KCND, which involved a number of strands and components, it was difficult to identify those that were likely to activate mechanisms leading to change outcomes. For instance, the various strands in the programme involved: the development of care pathways, an information campaign aimed at pregnant women, a leadership programme for consultant midwives along with a number of normality-related changes such as continuous risk assessment, midwife-led care for healthy pregnancies and discontinuation of CTGs on admission in labour. The authors worked to distinguish between those that represented the measures introduced to enable change and those that were activities designed to support the implementation of those measures. The former were referred to as programme components, each one of which will work through their own underlying mechanisms. The latter were classified as contexts as they were expected to facilitate the operation of the components. Accordingly, three programme components were identified: the appointment of consultant midwives in each health board, national pathways for maternity care and midwife as the first point of contact and lead care provider for healthy pregnant women.

Building the CMO hypotheses

The second challenge was in generating hypotheses about specific CMO configurations. Although the data were coded using the concept of contexts, mechanism and outcomes, the relationship among them was not self-evident. The task of breaking down the programme into its components helped the process of CMO generation, as an explanation could be built around each component in terms of how it would be expected to work, in what conditions and to what effects. For each component, data were drawn on that suggested the ways in which people were expected to respond to and act upon the changes brought on by the component (M1, M2, M3...), the factors anticipated to enable or constrain those mechanisms (C1, C2, C3...), and the consequences those mechanisms were expected to lead to (O1, O2, O3…). Using these data, a coherent CMO hypothesis was put together for each component. For example, the CMO in relation to the appointment of consultant midwives was as follows: the consultant midwives were expected to facilitate practice change through negotiation with all stakeholders, gaining multiprofessional engagement, acting as champions of normality, providing training and problem-solving. It was anticipated that the consultant midwives' experience, special interest in normal birth, and additional leadership training would increase their credibility and effectiveness

Abhyankar P, Cheyne H, Maxwell M, Harris F, McCourt C. (2013) A realist evaluation of a normal birth programme. *Evidence Based Midwifery* **11(4)**: 112-119

© 2013 The Royal College of Midwives. *Evidence Based Midwifery* **11(4)**: 112-119

as programme leads. However, hurdles were envisaged in engaging the multidisciplinary team and in potential role conflicts with senior midwife managers.

Stage two – testing the programme theory

The objective in stage two was to test the hypotheses that was developed in the first stage by collecting data in a range of different contexts, to explore the way in which the programme's anticipated mechanisms unfolded across different practice contexts.

Design

A multiple case study design was used as it enables in-depth study of a contemporary phenomenon within its real-life context (Yin, 2009) and complements the realist approach to evaluation. The size of the 'case unit' was determined by the need to capture the variation in process, context and outcomes of implementation at organisational and practice levels. Within each health board, maternity service provision comprises one or more maternity units varying in size, location, and model of care. To encompass the contextual conditions at a range of levels, a 'case' was defined as 'the maternity service in a particular health board area'.

Selection of cases

Three health board areas were purposively selected to maximise the opportunity for accessing a diverse range of contexts. To aid the selection, a sampling frame was first constructed using information from a health board level survey of maternity care practice prior to programme implementation. The parameters included in the sampling frame were demographics of the health board area (configuration of maternity services, annual births, demographic characteristics of the population and rurality) and the adoption of aspects of practice relevant to midwife-led care and normal birth (discontinuation of routine electronic monitoring on admission in labour and the midwife as the first point of professional contact in pregnancy). Information on the status of the health boards on these parameters was compiled in the form of a matrix. Boards with different combinations of parameters were selected to ensure diversity in case study profiles.

Data collection

Both qualitative and quantitative data were collected, using semi-structured interviews and focus groups with staff, and case record audit of all births in Scotland during one week, in two consecutive years before and after implementation of the programme (Cheyne et al, 2013). Within each case study site, the authors sought to interview personnel selected on the basis of their role in maternity services across levels: practice (frontline care staff) and organisational (staff from senior clinical and health board management). A purposive approach to sampling was used with the aim of selecting informants with a diverse range of views and experiences of KCND. At practice level, the authors planned to recruit clinical staff involved in maternity care within each case – at least two obstetricians, two GPs and two groups of midwives, hospital and community based, with five to seven per group. The sample included clinical leads relevant to maternity care and the health board management, specifically: head of midwifery, clinical director, director of nursing, KCND consultant midwife, and supervisor of midwives.

The interviews and focus groups centred on staff's views and experiences of being involved in the KCND implementation. The topic guides were informed by the realist framework to elicit information on three key elements:

- Context: views about the KCND initiative, the way programme components were implemented, the way change was facilitated, current practice and culture in the unit, the local context of maternity provision, and enabling and constraining factors
- Mechanisms: views of how the programme components worked, how the changes were interpreted and acted upon, and experiences of implementing the changes
- Outcomes: perceived changes in practice and service performance, impact on roles, workload and relationships.

Analysis and interpretation to develop middle-range theories

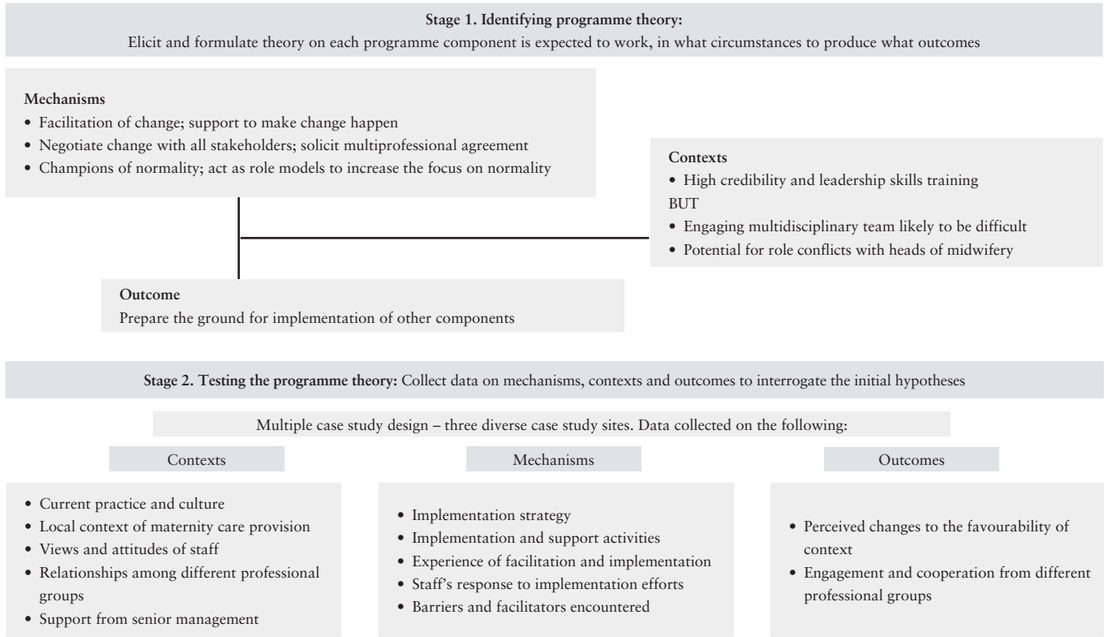
This phase aimed to analyse and interpret the data gathered in phase two, in order to understand the ways in which the proposed mechanisms unfolded in practice and identify alternative mechanisms and explanations. Data gathered during phase two was subjected to analysis using a process similar to that of phase one. Three transcripts of interviews – with a consultant midwife, a consultant obstetrician, a head of midwifery and one transcript of a focus group with clinical midwives – were selected to develop the initial coding framework. The transcripts were selected to ensure a fair representation of views across various organisational levels, as well as case study sites. Briefly, the analytical steps involved familiarisation with data, descriptive coding, grouping of codes using realist concepts, identifying a coding frame, applying the coding frame to all transcripts and adapting and refining the coding frame to allow new themes as they emerge. The final coding framework clearly reflected the categorisation of the data into contexts, mechanisms and outcomes for each of the three components. Once all the data had been coded, the following steps were taken to move from the themes and categories to refining the CMO models.

Developing case-specific CMOs

Firstly, a detailed narrative was developed for each case study site. Each site was described in-depth, in terms of the local context that characterised it. This involved not only the demographics of the population and the maternity service, but also the existing practice models, culture within the service, relationship between various professional groups and staff attitudes. Within each case, the authors then sought to test out the proposed CMO configurations in relation to each component using the data obtained during stage two. Specifically, the authors actively looked for evidence threads that suggested the different ways in which the proposed mechanisms of each component unfolded on the ground, the associated contexts and ensuing outcomes. For example, the initial CMO for the appointment of consultant midwives hypothesised that this would trigger a range of

Abhyankar P, Cheyne H, Maxwell M, Harris F, McCourt C. (2013) A realist evaluation of a normal birth programme. *Evidence Based Midwifery* **11(4)**: 112-119

Figure 1. Worked example of the appointment of consultant midwives



Stage 3a. Development of case specific CMOs

	Context	Mechanism	Outcomes
Case A	Medical model for care and culture of intervention but practice already changing in desired direction	Multiprofessional implementation group Regular staff updates, feedback and ongoing support, pathways launched and made highly visible	Staff buy-in mixed. Engaged pro-normality staff. Obstetricians perceived KCND as top-down midwife initiative
Case B	Less medical model of care – 'pro-normality'. Midwives work and make decisions quite independently. Most KCND policies already in places	A bottom-up approach to implementation. Subtle integration of KCND with local policy. KCND incorporated in CM wider role – avoid perception of KCND 'plant'	Obstetricians supportive but distant. No obvious requirement for change
Case C	Highly medicalised care model and culture of intervention. None of KCND policies were in place Staff's mindset – pro-intervention Unequal balance of power between midwives and obstetricians	Highly visible implementation activities. Multiprofessional group meetings, debates and discussion, visits to GPs. Tough, head-on implementation strategies. Non-cooperation from multiprofessional team	Resistance to change from midwives. Obstetricians – felt erosion of power and authority. GPs slow to come on board Perceived pressure on midwives to use pathways

Stage 3b. Cross-case synthesis to refine programme

Context	Mechanism	Outcomes
Medicalised culture, unequal power and authority between midwives and obstetricians	Resistance to change, feel top down, perceived erosion of power	Low engagement from stakeholders, difficulties facilitating, tough strategies
Pro-normality culture, midwives recognised as equals	Less resistance to change, stakeholders involved though distant	High engagement from stakeholders, smoother facilitation
Subtle implementation – pathways integrated and adapted to local guidelines	No obvious requirement for change	Greater use of pathways and adherence to KCND principles

implementation and support activities initiated and led by the consultant midwife to facilitate practice change within their local area. When analysing the data from case studies, the authors sought to identify what specific activities and efforts were undertaken, the way in which they were received and responded to by practice staff and what happened as a result. This was followed by a process of identifying and understanding the interactions between specific mechanisms, the contexts in which they were triggered and the associated outcomes. This enabled us to build an explanation for the fate of each component within each case study site. For instance, in site C, the consultant midwife planned and embarked upon a range of highly visible implementation activities, which soon had to be adapted to the needs of a highly medically-dominated culture (C1).

This context triggered her to adopt a series of tough, head-on implementation strategies, for example: frequent audits and monitoring (M1) and debates with medical professionals (M2), which resulted in feelings of undue pressure among midwives (O1) and perceived erosion of authority by the medical staff (O2), however, ultimately appeared effective in achieving important changes in certain practices, such as discontinuation of admission CTGs.

Actively looking for new, unanticipated mechanisms

The process of testing and refining the initial CMOs in each case study site also led to the identification of new, unanticipated mechanisms. Below one unanticipated mechanism that was triggered by the appointment of consultant midwives is illustrated. Midwifery leadership, in the form of consultant midwives at health board level, had been under discussion for some time and there was a growing inclination for their appointment in some health boards (cases B and C). The KCND programme provided a timely opportunity for materialising their existing plans and facilitated the process of appointing the consultant midwives. The health boards were given the freedom to decide whether and how these posts would be sustained beyond the duration of the programme. In all three case study sites, there was a strong buy-in to the programme from the board level, which was manifested through the keen interest in appointing the clinical leads in full-time, substantive posts. This formed a favourable context (C), triggering a new mechanism – termed a commitment mechanism (M) – which signalled to all the stakeholders that the health board was committed to driving the KCND initiative forward. The commitment mechanism provided a significant backing to the activities and strategies implemented by the consultant midwife (O).

Examining the interaction between components

Following the development of case-specific CMO models in relation to all the components, the ways in which the components interacted with each other, as well as with local contexts, were investigated. The authors actively looked for instances where the activation of mechanisms for one component depended on the outcomes from the implementation of another component. For instance, it became apparent that the successful engagement and buy-in from frontline staff (an anticipated outcome of the consultant midwives' activities) subsequently became the context for the implementation of care pathways and midwife as the lead professional components. In case B, the greater support and buy-in from all stakeholders paved the way for smoother implementation of the care pathways and resulted in greater adherence to the KCND principles. In case C, the resistance to change from the obstetricians affected the ability of midwives to adhere to the intrapartum care guidelines set out in the pathways, as obstetricians continued to dominate the intrapartum setting and intervene in low-risk pregnancies.

Developing refined CMOs

The process of building explanations in the form of CMOs specific to each case was followed by cross-case comparisons and synthesis. Each component compared and contrasted the CMO models emerging from the three sites. At the case-level analysis, the 'context' within the CMO models was closely tied to the individual case study site so that the case study site served as a proxy to all aspects of the context. However, at the stage of cross-case synthesis, the analysis was carried out at a higher level of abstraction and transcended the individual cases. The CMO models were refined by identifying the 'contextual factors' that were common across the cases and re-examining the associated mechanisms and outcomes. This meant that a particular CMO was now able to explain the workings of a component in more than one case site where the specific contextual factors were present. For example, a refined CMO stated that the appointment of consultant midwives worked by triggering a range of facilitation and support mechanisms, however, these mechanisms were only successful in eliciting engagement and buy-in from staff where the culture was more pro-normality, obstetricians were supportive and midwives were recognised as equals. This theory was borne out in all aspects of maternity service in site B, but was also supported in the antenatal care service in site A, which was characterised by this favourable context.

Discussion

The Scottish government KCND programme was a large scale complex programme of service change, implemented simultaneously in all health boards across NHS Scotland. The task of evaluating the programme posed a considerable challenge to the research team. This paper has described the way in which this challenge was addressed, the key methodological issues encountered, and, in particular, the theory-driven evaluation approach that allowed the authors to answer to some extent the question – was the KCND programme successful and, if so, how and why did it work?

The KCND programme was rolled out nationally across NHS Scotland and, therefore, an experimental research design was not feasible. The realist evaluation approach offered the opportunity to generate a theory for why, how and where an intervention or programme will work and to test this theory in a range of real-life contexts to examine why some elements worked well and how context appeared to shape the outcomes.

The method also enabled the authors to develop middle-range theories that helped in drawing some transferrable lessons about how and why programmes work. In line with realist thinking, the authors demonstrated that it is the programme mechanisms that are key to transferrable learning. While KCND programme components are likely to be maternity care specific, the mechanisms activated by the components in specific contexts may be transferrable to wider healthcare settings. For example, the appointment of consultant midwives was a component specific to the KCND programme and unlikely to be readily transferred to other kinds of healthcare or social programmes. However, the success of the commitment mechanism triggered by this component – signalling that the programme was supported and upheld at the top-level management – offers transferrable learning. Future programmes may build in strategies that could trigger such a commitment mechanism, which would in turn enhance the likelihood that changes are implemented. These mechanisms can be further tested in other programme evaluations, as well as used formatively when developing new programmes using CMO hypothesis to anticipate barriers and target resources, thereby avoiding the 'one size fits all' approach to programme development and implementation.

There were a number of limitations. At stage one, initial CMO configurations were developed and these were tested in stage two; however, the outcomes obtained were largely subjective. For example; there was a perceived erosion of authority among the obstetricians, and perceived increases in activities to support normal birth. Furthermore, the changes instigated by the KCND programme were directed at the level of service organisation and clinical practice, yet many of the anticipated outcomes would impact at the level of women receiving maternity care. It was not possible to undertake data collection from service users and this was a significant limitation. Finally, while the realist design allowed identification of what aspects of a programme worked, for whom and in what circumstances, it also gave rise to some uncertainties. For example, the authors were able to suggest the way in which a subtle implementation strategy worked in a context favourable to programme implementation (case B), but were not able to test how this strategy would have worked in a different context. This would have required a more experimental approach testing this implementation strategy in a different set of contexts. Hence the nature of the findings remains somewhat tentative, and this is characteristic of many realist evaluations.

Conclusion

Using realist evaluation enabled a theoretically informed and robust evaluation of a national programme of change in maternity care and the provision of information to policy-makers and key stakeholders at clinical practice level on the ways in which it may have worked to achieve its aims and areas that require further input. Transferrable lessons for development and implementation of other large scale programmes of change in the NHS and beyond were also drawn.

References

Blamey A, Mackenzie M. (2007) Theories of change and realistic evaluation: peas in a pod or apples and oranges? *Evaluation* **13(4)**: 439-55.

Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P. (2000) Framework for design and evaluation of complex interventions to improve health. *British Medical Journal* **321(7262)**: 694-6.

Cheyne H, Abhyankar P, McCourt C. (2013) Empowering change: realist evaluation of a Scottish government programme to support normal birth. *Midwifery* **29(10)**: 1110-21.

Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. (2008) *Developing and evaluating complex interventions: new guidance*. Medical Research Council: London.

Donabedian A. (1980) *Explorations in quality assessment and monitoring: the definition of quality and approaches to its assessment (volume one)*. Health Administration Press: Ann Arbor, MI.

Greenhalgh G, Robert G, Macfarlane F, Bate P, Kyriakidou O. (2004) Diffusion of innovations in service organisations: systematic review and recommendations. *The Milbank Quarterly* **82(4)**: 581-629.

Greenhalgh T, Humphrey C, Hughes J, MacFarlane F, Butler C, Pawson R. (2009) How do you modernise a health service? A realist evaluation of whole-scale transformation in London. *The Milbank Quarterly* **87(2)**: 391-416.

Ling T. (2012) Evaluating complex and unfolding interventions in real time. *Evaluation* **18(1)**: 79-91.

Mackenzie M, O'Donnell C, Halliday E, Sridharan S, Platt S. (2010) Evaluating complex interventions: one size does not fit all. *British Medical Journal* **340(c185)**: 401-3.

McEvoy P, Richards D. (2003) Critical realism: a way forward for evaluation research in nursing? *Journal of Advanced Nursing* **43(4)**: 411-20.

Medical Research Council. (2000) *A framework for development and evaluation of RCTs for complex interventions to improve health*. See: mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC003372 (accessed 3 October 2013).

Pawson R. (2002) Evidence-based policy: in search of a method. *Evaluation* **8(2)**: 157-81.

Pawson R, Tilley N. (1997) *Realistic evaluation*. Sage: London.

Pawson R, Tilley N. (2004) *Realist evaluation*. See: communitymatters.com.au/RE_chapter.pdf (accessed 3 October 2013).

Pawson R. (2013) *The science of evaluation – a realist manifesto*. Sage: London.

Richie J, Spencer L. (1994) *Qualitative data analysis for applied policy research*: In: Bryman A, Burgess R. (Eds.). *Analysing qualitative data*. Routledge: London: 173-94.

Robson C. (2011) *Real world research: a resource for users of social research methods in applied settings*. Wiley: Chichester.

Rossi PH, Lipsey MW, Freeman HE. (2004) *Evaluation: a systematic approach*. Sage, Thousand Oaks, CA.

Scriven M. (1991) *Evaluation thesaurus*. Sage: London.

Weiss CH. (1997) Theory-based evaluation: past, present, and future. *New Directions for Evaluation* **76**: 41-55.

Yin RK. (2009) *Case study research: design and methods (fourth edition)*. Sage: London.

Abhyankar P, Cheyne H, Maxwell M, Harris F, McCourt C. (2013) A realist evaluation of a normal birth programme. *Evidence Based Midwifery* **11(4)**: 112-119

© 2013 The Royal College of Midwives. *Evidence Based Midwifery* **11(4)**: 112-119