A realist evaluation of a normal birth programme

Purna Abhyankar1 PhD, MSc, Helen Cheyne2 PhD MSc, RM, RGN, Margaret Maxwell1 PhD, MA, Fiona Harris1 PhD, MA

1. Research fellow, Nursing Midwifery and Allied Health Professions Research Unit, University of Stirling, Stirling FK8 4AL Scotland. Email: purna.abhyankar@stir.ac.uk
2. Professor of mental health and deputy director, Nursing Midwifery and Allied Health Professions Research Unit, University of Stirling, Stirling FK8 4AL Scotland. Email: h.cheyne@stir.ac.uk
3. Professor of mental health and deputy director, Nursing Midwifery and Allied Health Professions Research Unit, University of Stirling, Stirling FK8 4AL Scotland. Email: m.maxwell@stir.ac.uk
4. Research fellow, Nursing Midwifery and Allied Health Professions Research Unit, University of Stirling, Stirling FK8 4AL Scotland. Email: fiona.harris1@stir.ac.uk
5. Professor of mental health and deputy director, Nursing Midwifery and Allied Health Professions Research Unit, University of Stirling, Stirling FK8 4AL Scotland. Email: m.maxwell@stir.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 7 March 2013.

Abstract

Background. Complex programmes are characterised by multiple components, acting independently and interdependently, at different levels and over a long period of time. Programme development, implementation and outcomes are therefore complex and often mixed patterns 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, and fail to acknowledge the active role played by the social systems into which programmes are implemented. (Ling, 2012.)

Research evaluation, drawn from applied social sciences, offers systematic procedures for understanding the systemic 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 observed changes in practice or not. Formative evaluation, on the other hand, 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 driven 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 is a taxonomic framework, among these, theory-driven evaluation, a similar framework 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 factors and contexts influence 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 credible, practical and valid.

Realist evaluation contends that it is not programmes that work, it is the people involved in the programme who make who it works. 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 understand 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: ‘the successful outcomes (how – causes – outcomes – O) only in so far as they introduce appropriate ideas and opportunities (mechanisms – M) to groups in the appropriate social and cultural contexts (contexts – C)’ (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. The second phase is the development of a context-mechanism-outcome (C–M–O) framework (or theory of change) which specifies the programme mechanisms (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 (C–M–O configuration) can explain the observed results in terms of differences in 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 act in similar but different contexts and in what conditions to bring about different outcomes.

This paper describes how the realist evaluation approach was used to evaluate the Midwifery Programme of Change in maternity care: the Keeping childhood 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.
The concept of normal birth as a ‘good thing’ has gained widespread acceptance in the UK and many developed countries. 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 programme was managed and monitored by agencies involved in quality and monitoring of health care programmes work in real-life situations.

The valuation therefore aimed to understand how the KCND programme was implemented, how the different components related to each other, and how they were expected to work, in what conditions and to what effects.

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 after 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, so 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 possible and the programme allowed for opportunistic comparison for the programme having multiple objectives and numerous work streams operating across different levels of the maternity care system, it brought together several interventions or practices of different types. It was a three-year programme with an existing evaluation framework – data segments containing discrete bits of information to be assigned to categories in subsequent analytical steps. Each unit was then assessed on its ‘normal’ or ‘deviant’ properties.

The valuation framework ( Ritchie and Spencer, 1994) was adapted to analyse the data generated from interviews and focus groups (Evaluation: Evidence Based Midwifery: 11(4):112-119). The key aspects of the realist evaluation framework – contexts, mechanisms and outcomes. These were selected to ensure that the mechanisms identified were realisable within the KCND 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 three transcripts were classified as 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. Using these data, a coherent CMO configuration was constructed. The codes were largely descriptive and merely reflected what 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 expected to work, in what conditions and to what effects. 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.

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 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 realistic 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 cases. 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, relationships 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 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.

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, site, 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. For each health board, maternity service provision comprises one or more maternity units varying in size, site, 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’.

The 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, site, 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. For each health board, maternity service provision comprises one or more maternity units varying in size, site, 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’.

The 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, site, 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 approach to sampling was used with the aim of selecting senior clinical and health board management). A purposive 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.
imputation 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 identified specific challenges 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, 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).

Active learning for new, unanticipated mechanisms

The process of testing and refining the initial CMOs in each case study site was now able to generate 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 from the board level, which resulted in 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 appointment of consultant midwives was a specific component of the KCND programme and unlikely to be readily transferred to other kinds of healthcare or social programmes. However, the site C, of the factors' that were common across the cases and re-examining the ‘context’ within the CMO models was closely linked to anticipated mechanisms. In site C, the consultant midwife planned and embarked upon the antenatal care service in site A, which was characterised in low-risk pregnancies.

Developing refined CMOs

The process of building explanations in the form of CMOs specific to each case study site. 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. 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 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 methodologies and strategies. 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 empirical research design was not feasible. The realistic evaluation approach offered the opportunity to combine a theory-driven evaluation approach with an implementation intervention or programme 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 transferable 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 transferable learning. While KCND programme components are likely to be maternity care specific, the mechanisms activated by the components in specific contexts may be transferable to wider contexts. For example, the appointment of consultant midwives was a specific component of the KCND programme and unlikely to be readily transferred to other kinds of healthcare or social programmes. However, the site C, of the factors' that were common across the cases and re-examining the ‘context’ within the CMO models was closely linked to anticipated mechanisms. 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.

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 policymakers 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. Transferable lessons for development and implementation of other large scale programmes of change in the NHS and beyond were also drawn.

References