



COMPLEXITY IN HEALTH

Consequences for research & evaluation

FEMHealth discussion paper

**Bruno Marchal, Sara Van Belle, Vincent De Brouwere,
Sophie Witter & Guy Kegels**

February 2014

Acknowledgments

The research relating to this paper has received funding from the European Union Seventh Framework Programme (FP7/2007-13) under grant agreement no 261449, under the FEMHealth project. FEMHealth is a research consortium including the following partners:

- University of Aberdeen (Impact), UK (coordinator)
- Agence de Formation, de Recherche et d'Expertise en Santé pour l'Afrique (AFRICSanté), Burkina Faso
- Centre de Recherche en Reproduction Humaine et en Démographie (CERRHUD), Benin
- Ecole National de Santé Publique, Morocco
- Institute of Tropical Medicine, Belgium
- Institut de Recherche en Sciences de la Santé (IRSS), Burkina Faso
- London School of Tropical Medicine, UK
- MARIKANI, Mali

For further details, please see: www.abdn.ac.uk/femhealth

Table of Contents

Summary	5
Introduction	6
Some key features of complexity theory	7
Simple, complicated and complex	7
Complex adaptive systems	8
Complexity in health	9
Complexity of health interventions	9
Complexity of health care organisations and systems	10
Complexity of policy making and implementation	11
Complexity & health policy and systems research	14
Five challenges	14
Dealing with complexity in health policy and systems research	16
A theory-driven approach to the FEMHealth project	21
FEMHealth's overall design	21
Using the programme theory as a structuring tool	22
Studying a complex problem - the challenges and some possible solutions	24
Conclusion	26

Summary

User fee exemption for delivery and emergency obstetric care (EmOC) is a policy that has recently been introduced by a large number of countries, particularly in Africa, with the aim of enhancing access to care and improving maternal and neonatal outcomes. In 2011, the FEMHealth project was established, with EC funding, to conduct multi-disciplinary evaluations of fee exemption policies in these four countries. FEMHealth started from the perspective that these are complex policies, requiring tailored evaluation methodologies. One of the project's objectives was therefore to develop new methodological approaches for the evaluation of complex interventions in low-income countries.

In this Discussion Paper, we explore the consequences of the notion of complexity for health policy research and evaluation through the lens of the FEMHealth research programme. We first present a definition of complexity and key elements of complex systems theory, applying these to policymaking and policy implementation. The frame of Zimmerman and Glouberman usefully differentiates 'complicated' and 'complex'. Key elements of complexity theory that are useful for health policy and systems research include non-linear interactions, feedback loops, time delays, path-dependence, self-organisation and emergent behaviour. We then apply these frames to health interventions, health care organisations and systems, policy making, decision-making and policy implementation.

If complex problems are characterised by unknown unknowns, uncertainty and emergence, there are 5 challenges for researchers and evaluators: capturing emergence, demonstrating causal attribution, answering the why question, learning from unique events, and accepting uncertainty.

In the health policy and systems research literature, two approaches can be found to deal with complexity: reducing complexity or embracing complexity. We present briefly examples of how complexity can be dealt with in a pragmatic way and how other approaches embrace complexity - including systems thinking and modelling and the theory-driven approaches. For the latter, we focus specifically on realist evaluation, an approach we used in some of the FEMHealth studies.

Finally, we describe how the FEMHealth project addressed the complexity of fee exemption policies. We used a key element of theory-driven evaluation, the programme theory, as a structuring tool to facilitate a structured discussion among the researchers of their own hypotheses. It also proved useful to coordinate work packages and review the programme's progress. We end with lessons learned in the process and some reflections on how to address complexity in health policy research and evaluation.

Thinking in terms of 'simple-complicated-complex' and trying to understand where a problem or situation can be located in the sense-making frames helps in choosing the most appropriate designs and research methods. Such frames also point out where and when 'hard' evidence can be obtained, and subsequently used in decision-making, and when we need to accept that evidence as construed by evidence-based medicine adepts can simply not be produced. In such cases, one can attempt to break down the complex issue into components, whereby we may need to accept that something will be lost in the simplification process. If that is not possible or desirable, adaptive approaches to research and evaluation are needed.

Perhaps the hardest thing is to see how in truly complex situations we can make sense of things, and to accept that complex issues can only be understood in retrospect. Only plausible explanations can be developed, not predictive theories. This does not mean that health policy and systems research reverts to journalism. Rather, building upon what is known and learning while doing becomes important. It is here that theory-driven inquiry approaches are most useful.

Introduction

User fee exemption for delivery and emergency obstetric care (EmOC) is a policy that has recently been introduced by a large number of countries, particularly in Africa, with the aim of enhancing access to care and improving maternal and neonatal outcomes (De Brouwere et al., 2010, Ridde, 2011).

The free caesarean section policy in Mali was introduced in 2005. It is applied nationally to all caesarean sections in the public sector, and in theory covers all facility-based costs (but not transport). In a three-way partition of costs, families are intended to fund the journey into the health centres, while communities fund the onward referral transport costs, and the state covers the costs of service provision, including accommodation, surgery, laboratory tests, and treatment of complications such as pre-eclampsia and ruptured uterus. Burkina Faso introduced a policy in 2006 that subsidised health facilities for 85 % of the cost incurred for normal deliveries and caesarean sections. This policy followed several other programmes introduced by the Ministry of Health to improve care for pregnant women. In Morocco, the fee exemption policy initiated in 2008 was comprehensive, abolishing all user fees related to prenatal clinic consultations, normal deliveries, caesarean sections and all required drugs and consumables. It was part of a broad action plan for the health sector, which also included a programme to improve supply of drugs, a health workforce plan and interventions aimed at improving transfers of patients between health facilities. In Benin, the policy introduced in 2009 was more selective, covering caesarean sections only and reimbursing health facilities with a flat fee for each intervention carried out.

In 2011, the FEMHealth project was established, with EC funding, to conduct multi-disciplinary evaluations of fee exemption policies in these four countries. A scan of the literature shows that the number of studies or evaluations of such policies is rising (Richard et al., 2010, McPake et al., 2011, Ridde and Morestin, 2011, Richard et al., 2013). These focus on policy effectiveness in terms of utilisation (Witter et al., 2010, De Allegri et al., 2011, Dhillon et al., 2012, Lagarde et al., 2012,), equity (El-Khoury et al., 2012) or cost-effectiveness (Witter et al., 2010). Others focus on implementation issues (Witter et al., 2007a, Ridde and Diarra, 2009, Nimpagaritse and Bertone, 2011, Ben Ameur et al., 2012, Witter et al., 2012, Idd et al., 2013) or barriers and facilitators (Ridde et al., 2010). Some studies focus on financing (Witter and Adjei, 2007), or assess the effects of such policies on the health workforce (Witter et al., 2007b, Carasso et al., 2012) or health facilities (Witter et al., 2011). Others still analyse the policy formulation process (Meessen et al., 2011, Witter et al., 2013). However, few of these studies are explicitly based on a hypothesis, a framework or a theory that would provide a basis for cross-case analysis or comparison of such policies. Exceptions include Walker and Gilson (2004), Gilson and McIntyre (2005), Ridde and Diarra (2009), Witter (2009), Hercot et al. (2011) and Robert et al. (2012).

FEMHealth¹ started from the perspective that these are complex policies. First, they aim to address high maternal mortality - a typically complex problem involving a large number of social, cultural, economic, personal and systemic factors - and second, their implementation is complex. This complex nature requires tailored evaluation methodologies. One of the project's objectives was therefore to develop new methodological approaches for the evaluation of complex interventions in low-income countries.

The importance of complexity for health care policy-making and implementation, as well as for research and evaluation, is now acknowledged (Gilson, 2012). However, in the policy and health systems research (HPSR) literature, conceptual confusion is reflected by the interchangeable use

¹ See www.abdn.ac.uk/femhealth for background on the project. The project ran from 2011 to 2014, and included partners from the UK (University of Aberdeen and the London School of Hygiene and Tropical Medicine), from Belgium (the Institute of Tropical Medicine), Burkina Faso (AfricSanté and IRSS), Benin (CERRHUD), Mali (MARIKANI), and Morocco (the National School of Public Health).

of terms such as ‘complicated’ and ‘complex’ and divergent definitions of what makes a problem, an intervention or a specific setting complex. Similar problems affect discussions on what constitutes good designs for evaluation or research of complex interventions (Kernick, 2006, Anderson, 2008,).

In this paper, we explore the consequences of the notion of complexity for health policy research and evaluation through the lens of the FEMHealth research programme. We first present a definition of complexity and key elements of complex systems theory, applying these to policymaking and policy implementation. We then identify the challenges of complexity for research and some approaches to deal with them. Finally, we describe how the FEMHealth project attempted to address the complexity of such policies. We end with lessons learned in the process and some reflections on how to address complexity in health policy research and evaluation (see also Marchal et al. (2013).

Some key features of complexity theory

The notion of complexity has its origins in the field of natural sciences. Complexity theory absorbed elements of general systems theory, cybernetics, chaos theory and information theory. In all these fields, an evolution from reductionist Newtonian models of a well-ordered universe to paradigms that focus on non-linear dynamics started in the 1950s. Later, complex systems thinking was applied in management, for instance by Stacey (1996a) and Zimmerman et al. (1998), and to the study of social phenomena by different social science disciplines (Byrne, 1998, Cilliers, 1998), to development (Ramalingam and Jones, 2008) and to policy analysis (Dennard et al., 2008, Swanson and Bhadwal, 2009).

In health, there was a wave of attention at the beginning of the millennium, calling for use of complexity concepts in health (Plsek, 2001, Plsek and Greenhalgh, 2001, Kernick, 2002, Kernick, 2004), and some focused specifically on complexity in management of clinical care (Priesmeyer et al., 1996, Plsek and Greenhalgh, 2001, Wilson and Holt, 2001). It took longer for complexity to surface in the mainstream of the public health literature. WHO, for instance, recently published a working paper on systems thinking and complexity in the frame of health system strengthening (de Savigny and Adam, 2009). This late adoption may be due, in part, to the conceptual confusion regarding the definition of ‘complexity’ and a fragmented application of complexity theory to health care (Begun et al., 2003, Rickles, 2008).

Before presenting some key concepts from complex adaptive systems theory, we first introduce the notions of ‘simple’, ‘complicated’ and ‘complex’.

Simple, complicated and complex

The distinction between simple, complicated and complex problems made by Glouberman and Zimmerman (2002) is a useful starting point. These authors relate their definitions to causality and solutions:

- **Simple problems** have simple causes. Causality is linear and simple problems have standard solutions. These can be applied without specific expertise; technical skills are sufficient.
- **Complicated problems** consist of sets of simple problems, but cannot be reduced to them. They are compounded by scale and coordination problems. Solving complicated problems requires expertise and collaboration between experts. Formulae and instructions to solve complicated problems can be developed and are critical to success. If experts apply the formulae correctly, outcomes can be predicted.

- **Complex problems** include sets of simple and complicated problems to which they are not reducible. The interactions between determinants of the sub-problems can lead to non-linear causal relations between potential causes and outcomes. Also context-sensitivity can make a problem complex. As a consequence, outcomes are unpredictable. To solve complex problems, formulae and standardised solutions that proved effective in the past provide little guidance. Instead, complex problems are solved through safe-fail experiments that allow learning by doing or by making sense of events post facto.

Financial barriers to utilisation of health care services and maternal mortality both fit Glouberman and Zimmerman's definition of complex problems, determined as they are by multiple, interlinked factors. A fee exemption policy for pregnant women, that in essence consists of abolishing user fees for a certain group of the population, may seem at first a simple intervention. It can be introduced by mere administrative fiat, targets a well-specified group and has a simple causal chain: abolishing user fees reduces financial access barriers and leads thus to higher utilisation by pregnant women. This in turn is expected to contribute to more timely case management of complications of pregnancy or delivery and ultimately to lower morbidity and mortality. However, the actual implementation and uptake of the policy, and thus its effect, depends on the actors involved. They are likely to adapt the policy to the local context. The policy outcome will also be influenced by pre-existing context factors and determinants like poverty levels, health system coverage, quality of care, etc.

Complex adaptive systems

The complexity of fee exemption policies can also be assessed using the terminology of complex adaptive systems theory. This requires us first to consider what a 'system' is. Morin (2001) defines a *system* as a unit made up by and organised through relations between elements (or agents), structures and actions (or processes). As with any system, *complex* systems consist of multiple elements, which interact with their environment, but some factors make them stand out: the nature of the interactions, the feedback loops, and the importance of the initial conditions and of the past. As a result, complex systems will display emergent behaviour and unpredictability. This applies as much to complex biological systems as to human social systems, including health systems (Plsek and Greenhalgh, 2001, Wilson and Holt, 2001).

To understand complex systems, one needs to understand the nature of the interactions between the elements. Typically, these interactions can be non-linear: small inputs may have large effects and vice versa. The effect of actions also depends on the initial conditions. In the case of a fee exemption policy, for instance, the result can be expected to be greater in regions with relatively high poverty levels compared to low poverty regions, assuming other barriers are similar.

In complex systems, positive and negative feedback loops contribute to emergent behaviour and unpredictability, and this is largely due to the human factor or the way human beings react to change. For instance, a policy that abolishes user fees may lead to higher utilisation of the hospital because it reduces financial barriers to access. This may lead to higher workloads for the health workers, and in response, health workers may impose new barriers to patient access in an effort to reduce stress. Other unintended effects may occur as overworked health workers become unfriendly to patients, leading to reduced patient satisfaction, which in turn may affect the decision to use the hospital's services. Such feedback loops can often explain unexpected or perverse results. Furthermore, feedback loops may display time delays, in which case effects only become apparent after long periods of time. If managers or policymakers overreact in response to slow results of an intervention by initiating new interventions, the situation can change wildly (oscillation).

Complex systems are also path-dependent: outcomes of interventions are sensitive not only to initial (current) conditions, but also to decisions taken in the past. Applied to policymaking, this explains how present policy choices and implementation modes are determined by past choices. Managers accustomed to raise organisational revenue by being paid fees for service by users, for example, find it hard to adjust to a fixed reimbursement per episode under exemption policies. This may explain why in Burkina Faso, for example, there has been a reversion to charging per item, contrary to the original official fee subsidy policy (Ministère de la Santé, 2006).

Some of the above features already hint at the ability to 'self-organise' that makes a complex system *adaptive*. Human agency is indeed the key factor that leads to adaptive change and evolution within complex systems. It also leads to variation in behaviour being the rule in complex adaptive systems rather than exceptional.

Applying the above concepts from complexity theory to user fee exemption policies, it could be argued that these policies offer an apparently simple solution (of changing the financing structure for specific priority services, thereby reducing financial barriers to utilisation) to the problem of high maternal mortality, a complex problem. The barriers to increasing service uptake are multiple, and changes to one factor are likely to lead to a ripple of reactions and feedback. Exemption policies rely on changing the behaviour of a wide range of actors, not least pregnant women and their households. Within the health system, multiple layers and organisations are involved. Furthermore, context and history play an important role, setting the scene and influencing the range of both policy and implementation options. The success of fee exemption policies is thus based on a large set of conditions or assumptions, which need to be made clear when developing a research design for policy analysis.

Complexity in health

We argue that health systems and organisations can be better understood by recognising their complexity – where relevant. However, notions and concepts of complexity theory are all too often transferred to other domains in a superficial way (Stacey et al. 2000) and little of the complexity debate has been taken up in research on policymaking and implementation in health. In this section, we introduce some key frameworks that can help in making sense of the complexity of decision-making and policymaking.

Complexity of health interventions

A first perspective is to consider the degree of complexity of interventions in health. There is not much debate on when an intervention is simple: simple interventions are addressing a clearly defined problem for a clearly defined group of people and work upon a single cause or mechanism of change (Sanderson, 2002).

There is far less agreement on what makes interventions complex. Judge and Bauld (2006) propose criteria such as the need for protracted negotiation on the goals and resource allocation, the length of the implementation chains and the need for integration of lessons learned from previous projects.

The guidelines of the UK Medical Research Council for research of complex interventions define such interventions on the basis of the following criteria (Craig et al., 2008a):

- the number of and interactions between components of the intervention
- the number of groups or organisational levels targeted by the intervention
- the degree of flexibility of the intervention that is permitted

- the behaviour required by those delivering or receiving the intervention
- the number and variability of the outcomes

However, this definition lacks some key features of complexity such as feedback loops, emergent behaviour and unpredictability (Anderson, 2008). It also considers quantity (i.e. number of interactions or target groups) as a factor of 'complexity', while this is perhaps better considered as a defining element of 'complication'. Finally, it does not pay much attention to the influence of the context, nor to the notion of co-development and path dependency. The guidelines are conflating 'complex' with 'complicated'.

A comprehensive definition of complex interventions

In the definition of a complex intervention, we would particularly stress the interaction between actors, context and intervention as making an intervention complex: it is the often unpredictable interaction between people and with their context that defines the actual outcomes. This reflects the definition of complex interventions of (Wong et al., 2010), who put human beings at the centre: *"Complex interventions consist of multiple human components that interact in a non-linear fashion to produce outcomes which are highly context dependent."*

The degree of complexity of interventions could then be defined in function of:

- the definition, scope and degree of 'wickedness' of the targeted problem
- the degree to which the problem is determined by multiple determinants (the problem context)
- the scope of the intervention (broad, ill defined targets)
- the level or locus of the intervention: structural complexity as a result of interconnectedness between levels (e.g. national-regional-local level of the health system; 'organisation-team-individual')
- the actors (number, interaction, interests, need for engagement, power)
- the expected interval between intervention and effect
- the mechanism through which the intervention is supposed to act (or the models of change maintained by the actors)
- the importance of history (path dependence)

Complexity of health care organisations and systems

Much of the complexity of implementing fee exemption policies stems from the nature of health systems. From an organisational theory perspective, health care organisations and systems present salient features of complex organisations (Trochim et al., 2006).

Any health care organisation consists of different units and layers. Each unit is dealing with different tasks and specialised functions, but they all need to be working towards the shared goal of providing services to patients and community. This would make them 'complicated' (Rickles et al., 2007). Health care organisations are made up by the people working in them and it is the central role of people in organisations that makes them 'complex'. Each staff member operates within the organisational goals and mission, but has personal aspirations and goals. Within organisations, people belong to specific cadres and professional groups, with their own specific competences, culture and shared mental models. These are elements that contribute to what Zimmerman et al., (2012) call 'structural complexity'.

However, each member of the organisation plays different roles: they are members of the organisation and of a professional cadre, but also of a family, a social group, a community, etc. Furthermore, people bring their 'whole self' to the organisation, not only their 'work self' (Linstead et al., 2009, p. 24). Unsurprisingly, members of the organisation do not always do what their job description prescribes as there are usually many sources of conflict of interests.

The diversity of people, their roles, identities and actions, as well as the tasks an organisation and its management team need to deliver, is reflected in different logics that are at work within health care organisations. Tensions exist between 'integration' and 'specialisation', but also between 'care and cure' and 'management and administration' (Glouberman and Mintzberg, 2001). Diverging interests occur not only at the level of individual providers, teams and units, but also at management level.

This can result in conflicts and turbulence, which managers normally consider as a problem for the organisation. Complexity theorists consider this also as a source of opportunities. New structures or behaviour may, indeed, result from self-organisation in response to the turbulence in the environment or as a consequence of the interaction between actors within the organisation. Such emergent behaviour, good or bad for the organisation, is essentially non-predictable.

However, health care organisations or systems and the people that make them up are not operating randomly. Health care organisations are path dependent: because of their history, decisions taken in the past may considerably restrain margins of freedom and future choices. Furthermore, the range of options open to systems – their response repertoire – may reduce over time as systems learn from experience what works best and adapt their processes and structure around these effective responses. Similarly, health care organisations are located in a web of institutional arrangements, social contracts, cultural contexts and political relations. Such elements of structure shape the trajectory of people and organisation – and at the same time are being shaped by the agency of people.

It is the human agency, the social relationships and interactions, the resulting feedback loops and the non-linearity of some of these interactions that make healthcare organisations socially complex (Glouberman and Zimmerman, 2002) and different from other complex systems in nature (Snowden and Boone, 2007).

Complexity of policy making and implementation

The notion of complexity has been taken up by several scholars in political sciences. Complexity thinking, indeed, provides useful perspectives both on the policymaking process and policy implementation, indicating how and why linear rational policy models are useful only for specific kinds of problems (Meek, 2010).

Complexity in policymaking

Many of the elements of the 'definition' of complexity that we introduced above are used to argue that the policymaking process is a complex process: the high number of actors who are involved at one stage or another, the long time-spans it may take to design and implement a policy, the multiple levels and programmes that need to be coordinated. Added to these factors of complexity are the technical and legal aspects and the conflicts of interests and values that often underlie policy decisions (Sabatier, 1999).

Although no comprehensive framework based on complexity thinking has been developed as of yet (Morçöl, 2010), a number of recently developed approaches to policy analysis incorporate aspects of complexity. These include the notion of multi-layered policymaking systems with multiple actors and organisations (Morçöl, 2010), the institutional analysis and development framework proposed by Ostrom (1990), the advocacy coalition framework of Sabatier and Jenkins-

Smith (1993) and network governance theories. From these schools, the following elements or insights emerge:

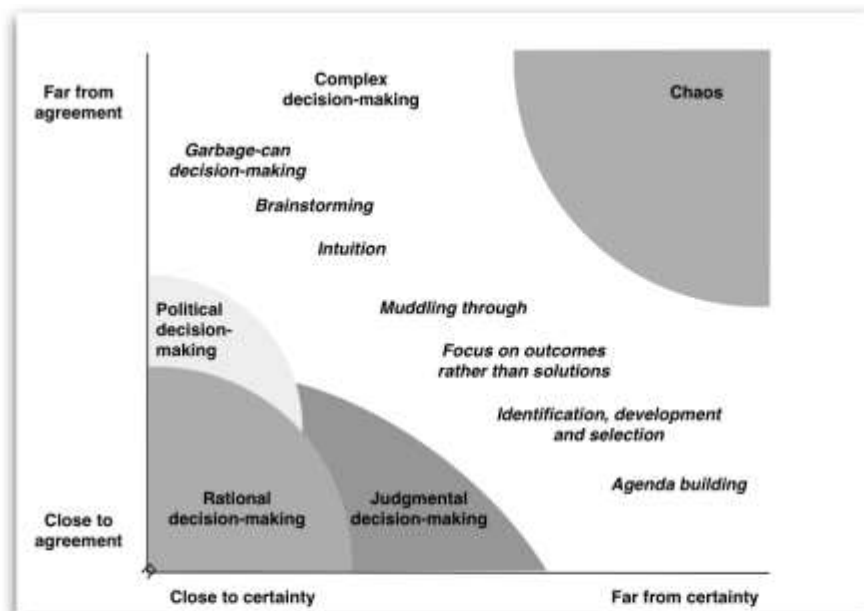
- Policymaking and development is intrinsically complex, to a large extent due to the multiple actors who are involved but also because of the complexity of the social and economic world (Boulton, 2010). “Policy itself is a self-organising and evolving system (not an external intervention into a self-organising system)” (Morçöl, 2010).
- Policies and the policymaking process are not only embedded in specific contexts, they are also dynamic because they co-evolve: “Public policies are self-organising systems that are constituted by the actions of self-conscious policy actors and they coevolve with other systems (natural systems and policy, social and economical systems)” (Morçöl, 2010).
- Policymakers can better deal with such complexity if they build in mechanisms for learning, continuous improvement and adaptation (Swanson and Bhadwal, 2009). Indeed, policymaking for complex issues can often be considered as an experiment from which systematic learning should ensue by incorporating learning and evaluation from diverse actor perspectives (Boulton, 2010).

These elements resonate well with the approaches to decision-making in the complex zone of Stacey’s diagramme (see below).

Complexity in decision-making

A framework that has been influential in a large number of domains is that of Stacey (1996b). The Stacey diagramme allows a choice of approach to decision-making in relation to different levels of agreement and certainty.²

Figure 1 - The Stacey diagramme



² Stacey distanced himself somehow from his diagramme as he moved towards stressing the central role of human relations at local level in creating outcomes (see Kernick, 2004)

This diagramme is useful in the analysis of decisions from policymaking over programme design and implementation to decision-making at the operational level of the health system.

- For issues that are close to certainty in terms of the link between action and outcome and close to agreement among the involved actors concerning the relation between action and outcome, rational decision-making models are effective. This zone contains simple issues for which best practices exist or for which one can learn from the past.
- Political decision-making is the mode for issues for which it is well known how to reach the outcomes (close to certainty) but for which there is little agreement on which outcomes should be prioritised. The decision-making process involves negotiation and compromise, and decisions tend to be made by the dominant coalition.
- Judgmental decision-making applies to problems for which agreement is high but certainty as to the cause-effect links low. A strongly shared vision may help in leading the organisation.
- The chaos zone is determined by high levels of disagreement and uncertainty, leading to anarchy. Leaders should move the organisation as fast as possible out of such zones through a command and control approach.
- In between the previous zones lies the complex decision-making zone, in which methods such as garbage-can decision-making (March and Olsen, 1976, Dixon-Woods et al., 2004), muddling through and incremental decision-making (Lindblom, 1980) and agenda-setting (Kingdon, 1995) are effective approaches to decision-making (Figure 1).

It should be noted that Zimmerman et al. (undated) presents a simplified diagramme, whereby the rational decision-making zone is called the simple zone, the political and judgmental zones are combined in the complicated zone, and complex decision-making falls in the complex zone.

Complexity in policy implementation

The complexity perspective has been applied to the problem of policy implementation by Jones (2011), who assesses the degree of complexity of the problems that policies can tackle according to three categories.

- The first category is the knowledge on cause-and-effect and on the means for addressing the issue: is it well established or not?
- The second category concerns the degree of consensus on the issues that a policy needs to address or the goals it needs to achieve: is there consensus or dispute? These two categories can also be found in the Cynefin framework (Kurtz and Snowden, 2003, Snowden and Boone, 2007) and Stacey's diagramme.
- Jones adds the capacity to act as a third category: can the problem be tackled through a policy that is carried out by a hierarchically organised programme, or does it require a distributed capacity?

From this, he argues, three other questions ensue: (1) *Where* does the decision-making take place? (2) *When* is relevant knowledge about the action gained and when do important decisions need to be made? and (3) *How* can the decision-making process be made as effective as possible?

In this view, policy problems are likely to be complex if:

- the knowledge on cause-effect is limited (and thus predictability of outcomes is low)
- the consensus on policy issues and goals is limited (and thus divergence of actors' goals and the subsequent chance of conflict is high)

- the required capacity to implement the policy is distributed (and thus requiring intensive communication and negotiation with many actors at all levels of the health system)

In practice, implementation of complex policies is likely to fail when inappropriate top-down approaches are applied. A 'best practices approach' is well suited for well-structured problems, or simple contexts, while a technocratic expert-based approach is good for moderately structured problems or complicated contexts (Snowden and Boone, 2007). However, because of their essentially linear assumptions about both the causality of policy problems and the implementation of policies, such approaches do not fit complex policies and contexts.

Complexity & health policy and systems research

Five challenges

The above discussion of complexity theory and its application in a number of analytical frameworks and fields clearly shows that health and health problems, as well as interventions, health organisations and health systems, often have complex aspects. Research on policies will be need to deal with the following factors of complexity:

- The presence of multi-faceted and long causal chains
- Significant time lags between policy decision, implementation and outcomes and the consequent risk of mismatch between research and policy time frames
- Tracking policies that in general will not go according to plan and that will lead to unexpected outcomes
- Assessing the co-evolution of the policy with other policies and key context elements
- Attributing the observed outcomes to the policy in question
- Accounting for the influence of context on the policymaking process, the implementation process and the outcome.

This complex nature of policies (and of health systems in general) leads to five challenges.

Capturing emergence

If unknown unknowns, uncertainty and emergence characterise the complex zone, research and evaluation designs should be able to deal with that. Designs should allow picking up alterations of the planned intervention, as well as parallel events or context elements that might affect the implementation and/or outcomes. More importantly, designs should deal with the social interaction that leads to emergent behaviour (see also the next challenge). To do so, they need to be flexible, adaptive and aim at learning (McDaniel et al., 2009). Using fixed protocols that are based on the assumption of predictability is an example of using a strategy for complicated issues in the wrong zone. Yet, in most cases, funders and ethical review boards demand detailed research protocols with as little of uncertainty and flexibility as possible.

The impossibility of identifying the root causes of a complex problem or of predicting when and how an intervention will move to the complex zone does not mean that researchers cannot assess the range of reasonable options: their capacity for anticipation will be enhanced if a wide range of observation and collection methods is used. Longitudinal approaches such as processual analysis, a method developed by (Pettigrew, 1990, Pettigrew, 1997) are well suited to capture unfolding interventions and the response of actors in the light of a configuration of other interventions and close and distal context elements. The HPSR Methodological Reader shows how a wide variety of

studies used a range of approaches to deal with multiple perspectives and complex causality (Gilson, 2012)

Demonstrating causal attribution

A fundamental challenge in HPSR concerns the assessment of the effectiveness of complex interventions and of causal attribution. Complex problems are by definition multi-determined, and interventions targeting such problems therefore intervene in complex webs of causality, easily triggering non-linear effects through negative and positive feedback. This makes it difficult to assess the relative contribution of the intervention to the observed outcome. Demonstrating that the intervention is necessary to cause the outcome is therefore only possible if the issue is simple or complicated. In other words, if probabilistic evidence can be produced, the issue is no longer complex. This reflects the notion that once the unknown unknowns can be known – and research is all about this – the issue has been moved out of the complex zone. However, complex problems can only be understood *a posteriori*, and the best researchers and evaluators can achieve is *ex post*, plausible explanations. At the same time, “*hindsight does not lead to foresight because external conditions and systems constantly change*” (Snowden and Boone, 2007).

Answering the why question

As difficult as the attribution question, and intricately related to it, is the question of identifying the mechanisms that explain change. The notion of mechanism can be interpreted in different ways. Within the realist evaluation approach (Pawson and Tilley, 1997), mechanisms are at the heart of the explanation: it is not interventions, but people who change problem situations. Interventions are resources taken up (or not) by people and contribute to the outcome only if they trigger mechanisms of change (see below). Such mechanisms can be situated at different levels: within the individuals and teams that make up the organisation, at the organisational level, or within society. Mechanisms are thus psychological, social or cultural in nature. Mechanisms of change are likely to be multiple, interacting with each other and ‘firing’ differently in different contexts. Only by understanding the mechanisms can we provide plausible explanations and useful advice to decision-makers who operate in other contexts.

Learning from unique events

Another challenge is to learn as much as we can from meagre data (McDaniel et al., 2009). Typically in complex systems, emerging events are unique (to a certain degree) and this explains the interest in better methods for single-case studies or ‘small n’ studies (see for instance White and Phillips, 2012). More methodological work, however, is needed to deal with this challenge.

Accepting uncertainty

If complexity is in essence about dealing with unknown unknowns, uncertainty and emergence, research on and evaluation of complex problems calls for particular attention to researchers’ mind sets. Research of complex issues requires “*a willingness to be uncertain at times and to know that being uncertain is crucial to the process*” (Zimmerman et al., 2012). This calls for reflexive researchers, who know when they cannot really know what is going on, and who are capable of double loop learning. Reflexive practice (Argyris and Schön, 1978, Schön, 1983), Kolb’s experiential learning (Kolb, 1984) and learning organisation theory (Senge, 1990, Garvin, 1993) are examples of approaches that help expert researchers to decontextualize their experience and recontextualise their knowledge and know-how – moving from single loop to double loop learning. The existence of unknown unknowns also means that health policy and systems researchers need to realise that expertise is relative when dealing with truly complex issues. This is no easy feat, as expertise is highly valued both in academic and evaluation circles.

Dealing with complexity in health policy and systems research

While the challenges may be increasingly acknowledged, there is far less agreement on how researchers and evaluators should tackle complexity. The literature shows that there are basically two main approaches: some scholars and practitioners call for reducing complexity, while others embrace complexity.

Reducing complexity

Forss et al. (2011) provides four principles for dealing with the challenges of evaluating complex interventions:

- *Be concrete*: understand the level of complexity and choose an appropriate evaluation design. Indeed, not all evaluations need to embrace a complexity perspective. This is reflected in the approach of Rogers (2011), who presents ways to assess the degree of complexity of an issue. However, if the complex approach seems the most appropriate, then methods should follow. Forss and colleagues advocate for the use of models to illustrate complexity, manage complexity and establish criteria of assessment and assess outcomes, but as we will see below, alternative approaches exist.
- *Be inventive*: deal innovatively with the causal attribution issue by using a multi-faceted approach. The authors refer to contribution analysis (Mayne, 2001, Mayne, 2011, Mayne, 2012) and intervention path contribution analysis (Schwartz and Garcia, 2011) as useful approaches to draft and test programme theories that explore and demonstrate causal contribution.
- *Be flexible*: Forss and colleagues recommend adopting multiple perspectives and multiple theories of change.
- *Be specific*: while flexibility is important, researchers and evaluators still should respond with tailor-made and specific responses to the evaluation task.

Some researchers have come around and developed strategies to deal with the real-life complexity of health systems that are not method-driven. Victora et al. (2011), for instance, argue that much of the current debate on complexity in health policy and systems research is focusing on evaluation of large-scale interventions (see Craig et al., 2008b). Instead of starting from methodological preferences and exigencies, Victora and colleagues propose to take the operational level as starting point. The difficulty of capturing the actual implementation of large scale interventions in different settings in one unique design and the difficulty of using comparison between intervention and non-intervention cases as the analytical device leads these authors to advocate for building such evaluations around the health district as the unit of analysis. The authors call for developing national research platforms centred on the district level, with continuous monitoring of health care facilities, programmes and services using a wide range of relevant indicators.

Pragmatic approaches to complexity

Some authors in the discipline of evaluation present ways to reduce the complexity of the issue at hand to make it more manageable. Rogers, for instance, proposes guiding principles for describing programmes or interventions in terms of 'simple/complicated/complex', which then can help to choose appropriate evaluation methods (Rogers, 2009, Rogers, 2011). Drawing upon the distinctions made by Glouberman and Zimmerman (2002), she proposes the following criteria to describe interventions: focus, governance, consistency of the intervention, 'necessariness' of the intervention, sufficiency of the intervention, the change trajectory (causal chain), and unintended outcomes (Table 1).

Rogers considers this framework to be a heuristic or a tool to think about interventions and how to evaluate them and argues that it could help to choose the most appropriate design and to avoid doing overly complicated evaluations when not necessary. In case of complex issues, she refers to Patton's developmental evaluation and to building and refining recursive programme theories (see below).

Table 1 - Criteria for choosing evaluation designs (Rogers, 2009)

Dimension	Simple intervention	Complicated intervention	Complex intervention
Focus	Single set of objectives	Multiple competing objectives maintained by different actors	Emergent objectives, defined in the course of the intervention through negotiation
Governance	Single organisation responsible or involved in steering the intervention	Multiple organisations	Emergent organisations (e.g. partnerships)
Consistency of the intervention	Application standardised through protocols	Body of knowledge adapted by experts	Adaptive approach is required
Necessariness of the intervention	Intervention is the only way	Intervention is one of several possible ways and can be identified as such beforehand	Intervention is one of several possible ways and its contribution can only be identified post facto
Sufficiency of the intervention	The intervention is sufficient in itself to produce the intended impact	Intervention requires other elements; effect of intervention is partial and can be foreseen	Intervention requires other elements, effect of intervention is partial and only be understood post facto
Change trajectory (causal chain)	Change trajectory is simple and readily understood	Change trajectory needs expertise to be understood	Change trajectory can only be understood in retrospect
Unintended outcomes	Unintended outcomes can be readily anticipated	Unintended outcomes can be anticipated by experts	Unintended outcomes cannot be anticipated but only identified as they emerge

A pragmatic approach to policy analysis

Adherents of the pragmatic school argue that methods and approaches for policy analysis and research are best built into the policy implementation itself.

- If the evaluation question is about effectiveness and impact, constituting a counterfactual is important in the process of assessing the actual contribution of the intervention. Usually, the counterfactual – what would have happened in the absence of the policy – is established by having control cases.
- If true experiments cannot be carried out – as most often will be the case in policy evaluations (Milton et al., 2011) - the policy implementation process can sometimes be modified to provide evaluation opportunities. This includes pilots, often used to test at small scale whether and how a policy works, randomisation of intervention sites (RCT design), phased introduction (sequential exposure of all participants over time) and intermittent operation (short term exposure in bursts). Identifying and using objective allocation rules may help evaluation by providing criteria through which participants, and thus exposure or coverage can be differentiated (HM Treasury, 2011).

- If the implementation process cannot be modified, universally implemented policies are best considered as natural policy experiments.
- If outcomes need to be assessed, observational outcome evaluation designs such as interrupted time-series or before-and-after designs can be used (Milton et al., 2011, Lagarde, 2012).
- If the emphasis lies on implementation issues, process evaluation, action research and case studies can be used. The advantage of such approaches is mainly that they focus by their very nature on the implementation phase and that they allow for identifying and assessing unpredicted emergent behaviour and outcomes. Action research has the additional advantage that it is in essence an adaptive method that helps to learn from actual practice in a systematic manner.

Embracing (social) complexity

The other main school argues that if *social* complexity is a key issue that is at present relatively neglected on the methodological front, methods need to focus more on relations between people, organisations, communities and other system parts. They need to be able to deal with the key issues of loosely-coupled networks of actors that make up the health system (and thus non-linear relations, time delays in feedback loops, self-organisation and emergence of new behaviour), the embeddedness of health systems in multi-layered contexts and systems and the resulting co-evolution that occurs permanently. Research and evaluation approaches need to provide a holistic and systemic view on the problem and/or solution. This is the challenge of scope, and not many methods are well suited for this (Sterman, 2006).

In health policy and systems research, several approaches are currently being used, including systems thinking and the related modelling approaches (see below). In this literature, there is less attention for some interesting approaches that have been developed in other disciplines, notably in evaluation. Patton's developmental evaluation approach, for instance, calls for continuous adaptation of the evaluation to the evolving intervention and to changing contexts through use of emergent evaluation designs and techniques that allow for participative engagement of the stakeholders (Patton, 2011, Rogers, 2011). Another useful approach is realist evaluation (RE), which was developed by Pawson and Tilley (1997). RE is part of the school of theory-driven inquiry and below we discuss the main schools and their relevance for research of complex issues.

Systems thinking

During the last ten or so years, 'systems thinking' emerged as a term often used by advocates of new approaches to complexity for research, intervention design, development and evaluation in public health. The American Journal of Public Health dedicated almost an entire issue to it in March 2006, while WHO and the Alliance for Health Policy and Systems Research published a flagship report on systems thinking for health system strengthening in 2009 (de Savigny and Adam, 2009). 'Systems thinking' seems most often used in the definition given by (Leischow and Milstein, 2006): "*a paradigm or perspective that considers connections among different components, plans for the implications of their interaction, and requires transdisciplinary thinking as well as active engagement of those who have a stake in the outcome to govern the course of action.*" This reflects some key notions of complexity. Its advocates argue that all too often, in problem framing and analysis, policymaking and evaluation the features of complexity are ignored, resulting in fragmented and ineffective policies and programmes.

Systems modelling is related to systems thinking and is perhaps best considered as a methodological tradition that aims at using modelling techniques to help understand complex systems (Trochim et al., 2006, Garnett et al., 2011). Examples of studies using systems modelling include Ng et al. (2011) and Chandrasekaran et al. (2008). Drawbacks include the fact that models require precise descriptions of the assumptions about the parameters and processes that make up

the causal web. This may be applicable to disease transmission or disease development (if the models apply non-linear pathways), but in case of human systems, such assumptions are much harder to describe.

Modelling techniques from the domain of *systems dynamics*, developed by Forrester in the 1950s, have been applied to deal with the dynamical complexity of systems. System dynamics holds that accumulation of people, assets, information, etc. and the resulting feedback mechanisms lead to complex behaviour of organisations. Feedback loops, stock and flow and time delays are typically elements of system dynamics models. This method has been used in epidemiological studies of diseases and substance abuse, in and in health care organisation studies in order to explore how feedback influences system elements and compare scenarios for policymaking (Homer and Hirsch, 2006, Midgley, 2006).

Agent-based modelling is a simulation approach based on individual entities (agents) and behavioural rules that govern their interaction. It allows considering inter-individual variability (Resnicow and Page, 2008).

Theory-driven evaluation

Theory-driven evaluation³ emerged in the discipline of evaluation during the 1980s, growing out of the programme theory field (Dickinson, 2006). Chen and Rossi developed it as an answer to policy and programme evaluation approaches that remained limited to before-after and input-output designs or that focused narrowly on methodological issues (method-driven evaluation) (Chen and Rossi, 1980, Chen and Rossi, 1983, Chen and Rossi, 1987).

Chen and Rossi argued that for any intervention, a programme theory could be described that explains how the planners expect the intervention to reach its objective (Chen, 1989, Chen and Rossi, 1989). Describing the often implicit set of assumptions that steer the choice and design of a programme or intervention is useful, because it allows an understanding of what is being implemented and why.⁴ Theory is defined by Chen and Rossi (1983) as the “*prosaic theories that are concerned with how human organizations work and how social problems are generated*”. Programme theories represent models or hypotheses that can be tested and further developed.

While theoretical publications on theory-driven evaluation appeared during the 1980s, there were few reports of empirical studies (Weiss, 1997b). It emerged again at the European Conference of Evaluation in 2002 (Van der Knaap, 2004). Methodological developments had continued, however, in the field of programme evaluation by authors like Chen (2005) and Donaldson (2007). The latter author developed the *program* theory-driven evaluation approach, which aimed at updating theory-driven evaluation and making it more practical. Applications in health include Sidani et al. (2004), Judge and Bauld (2006) and Perkins et al. (2007) and in health policy and systems research in the South, Van Belle et al. (2010), Nabyonga Orem et al. (2012) and Nabyonga Orem et al. (2013).

Theory of change

The theory of change (TOC) approach⁵ was developed by the Roundtable on Community Change (Aspen Institute) to evaluate complex community-based change interventions (Connell et al., 1995, Weiss, 1995, Fulbright-Anderson et al., 1998). More pragmatic in approach and oriented towards stimulating practical change, TOC was applied in evaluations of community-based programmes that typically involved many agencies and actors, had several levels and strands of activities, objectives and strategies that shift in time and outcomes that are difficult to measure (Judge and Bauld, 2001). TOC is essentially prospective (Dickinson, 2006) and seeks to establish the links

³ Theory-driven evaluation or theory-based evaluation: in the literature, these terms are often used interchangeably.

⁴ Sturmberg & Martin (2009) use the term ‘frames of understanding’.

⁵ The terms ‘theories of change’ and ‘theory of change’ are often used interchangeably.

between intervention, context and outcome (Weiss, 1995, Barnes et al., 2003, Mason and Barnes, 2007) through development and testing of logic models (Douglas et al., 2010). Such logic models describe the populations that are targeted, the indicators used to monitor change, the threshold indicators of significant change and the time lines (Judge and Bauld, 2001). Connell and Kubitsch (1998) argue that TOC can help in clarifying the attribution issue that hangs over so many complex interventions by its capacity to spell out the hypothetical causal chain of change and to test this.

In the USA, TOC has been integrated in programmes from the start, while its application in the UK was most often in evaluations of programmes that had already started (Mason and Barnes, 2007). Examples of studies that applied TOC in health care include Ying Ho (1999), MacKenzie and Blamey (2005), and Mason and Barnes (2007).

Realist evaluation

Realists argue that much can be learned in complex situations if the middle range theory is used as a starting point that gives structure to the inquiry.

Realist evaluation shares an emphasis on theory with the other schools of theory-driven inquiry: realist evaluations start with a theory and end with a theory. Theory should in this case be understood as middle-range theories as defined by Merton (1968, p. 39): “*theories that lie between the minor but necessary working hypotheses (...) and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behavior, social organization and social change*”.⁶

Realists hold that research and evaluation of complex issues cannot produce universally applicable findings. However, it can formulate plausible explanations that indicate in which specific conditions a particular programme works (or not) and how. Going from one case to the next, the middle range theory is ‘specified’ in a process of cumulative testing. It is this commitment to build and test middle range theories that gives realist evaluation the edge in complex evaluations, compared to more pragmatic approaches, such as Patton’s developmental evaluation approach (Patton, 2011), or even theory of change (Connell et al., 1995, Connell and Kubitsch, 1998).

It is the realist view on causality that makes realist evaluation well suited for the study of (social) complexity (Reed and Harvey, 1992, Westhorp, 2012). In short, this holds that actors have a potential for effectuating change by their very nature. Accepting as such the role of actors in change (agency), realist evaluation considers structural and institutional features to exist independently of the actors and researchers. Both actors and programmes are considered to be rooted in a stratified social reality, which results from an interplay between individuals and institutions, each with their own interest and objectives. If all human action is embedded within such a wider range of social processes, then causal mechanisms reside in social relations and context as much as in individuals. Accepting that not programmes but people change situations or solve problems, realist evaluation by definition deals with social complexity.

In practice, the realist method aims at identifying multiple causal pathways by using the Context-Mechanism-Outcome configuration as an analytical tool. Repeated studies lead to accumulation of insights in terms of CMO configurations, which can be translated in typologies of interventions and contexts. This may help to reduce the causal attribution problem (MacKenzie and Blamey, 2005, Leone, 2008).

While a number of challenges concerning the application of realist evaluation remain (Marchal et al., 2012), realist evaluation is increasingly used in health policy and systems research. Many such studies were carried out in the north (see for instance Byng et al., 2005, Tolson et al., 2007, Byng

⁶ Lipsey (1993) defines middle range theories as small theories or theoretical frameworks on which a meaningful differentiation of input, causal process and output can be based and which can help to unpack the black box between intervention and outcome.

et al., 2008, Leone, 2008, Greenhalgh et al., 2009, Mackenzie et al., 2009, Manzano-Santaella, 2011), but increasingly studies carried out in health in low and middle income countries are being published (Blaise and Kegels, 2004, Marchal et al., 2010a, Marchal et al., 2010b, Maluka et al., 2011, Prashanth et al., 2012, Robert et al., 2012, Goicolea et al., 2012,).

Pawson applied the realist approach to review and synthesis of evidence for policymakers, which he called realist synthesis (Pawson, 2002, Pawson et al., 2005, Pawson, 2006). Examples in health policy and systems research include (Dieleman et al., 2009, Robert et al., 2010, Kane et al., 2010).

A theory-driven approach to the FEMHealth project

FEMHealth's overall design

One of FEMHealth's objectives was to improve the knowledge base regarding the effectiveness, cost and impact of the removal of user fees for delivery care by carrying out comprehensive evaluations. The above discussion of complexity points to a number of consequences for the choice of research design and methods. Ideally, a study design for a complex intervention or problem should allow researchers to assess not only effectiveness but also the underlying processes so as to uncover the causal mechanisms. An understanding of how and in which context such policy can be expected to have similar impacts is central to its transferability. To do this, the study should explore the influence of key actors (including power analysis), and assess the organisational, social and historical context as well as the evolution of other policies that might have had an effect on the policy making process, implementation and observed outcomes. The design should deal with the significant time lags between policy decision, implementation and outcomes and the consequent risk of mismatch between research and policy time frames. Perhaps most challengingly, the design should be adaptive and allow for capturing the unexpected.

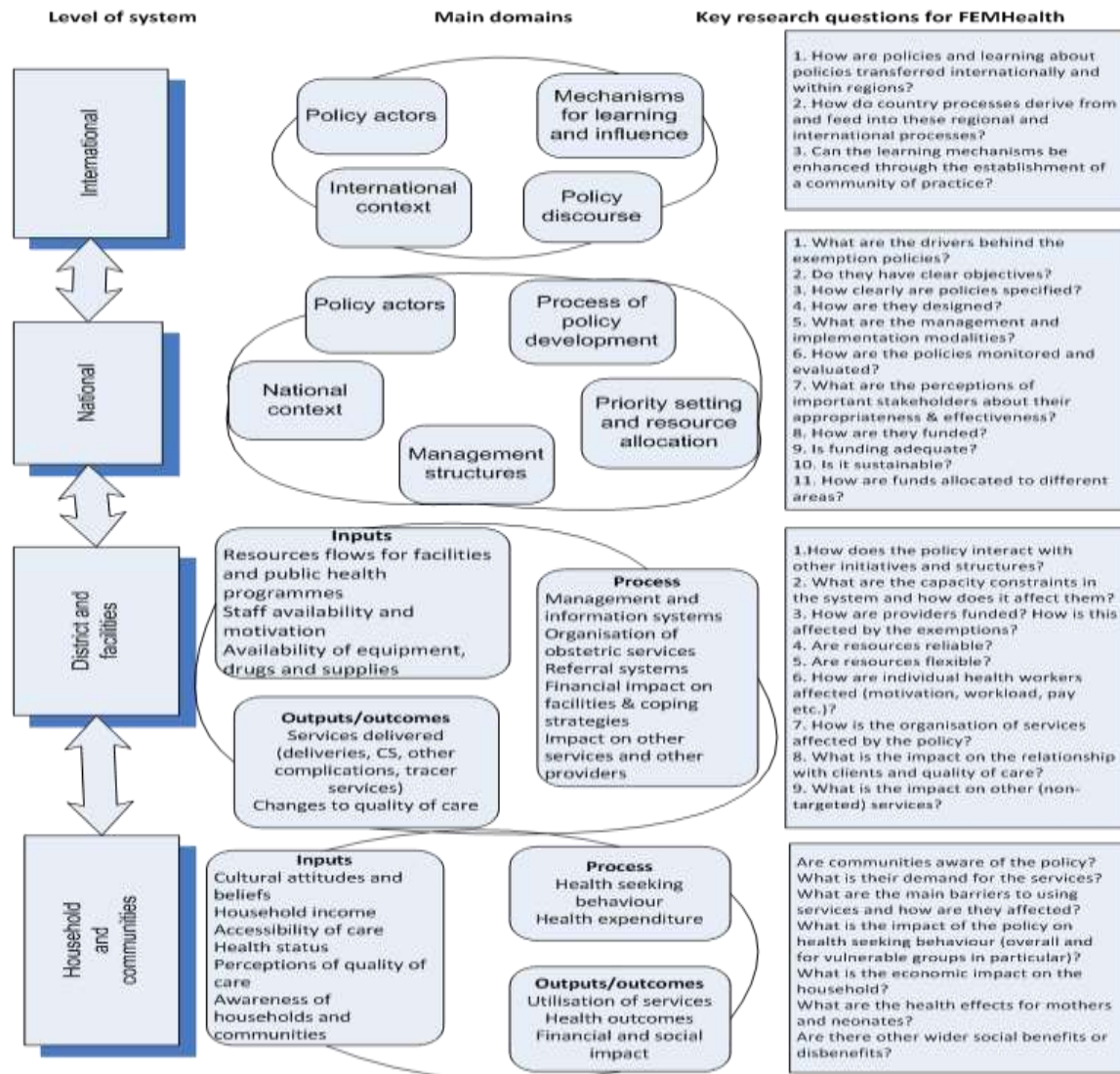
To deal with these challenges, FEMHealth adopted a multi-country, comparative case study design within a natural policy experiment perspective (Milton et al., 2011). In principle, case studies allow for a holistic in-depth investigation of issues as they happen in their natural setting, whereby different sources of information and data collection methods can be used concurrently (Denscombe, 2003). The case study design is in essence an adaptive design, as it facilitates exploration of a "phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" (Yin, 2003).

In practice, we selected between 6 and 8 study sites in each country. However, a series of case studies of districts or facilities aimed at studying the implementation process would be insufficient. We also focused downstream and set out to assess quality of care and other outcomes at patient level. Furthermore, political sciences studies showed how the policy formulation and translation into a programme influence the actual implementation of a policy (Berman, 1978, Sabatier and Mazmanian, 1979). For this reason, the policy formulation process and the arrangements put in place by the central level were examined as well, alongside some investigation into the interaction with regional and international ideas and actors. We thus aimed at covering the multiple interactions between the spheres of communities and pregnant women, service providers, service managers, programme managers and policymakers. In practice, we combined assessments of the policy formulation, the implementation processes, provider and user perspectives, the intermediate outputs and the outcomes with qualitative and quantitative methods and tools.

Using the programme theory as a structuring tool

Figure 2 presents how, during the preparatory phase, the different elements of the FEMHealth programme were conceived. It indicates the main domains of investigation and research questions for each level of the health system.

Figure 2 – The initial conceptual framework of the FEMHealth programme

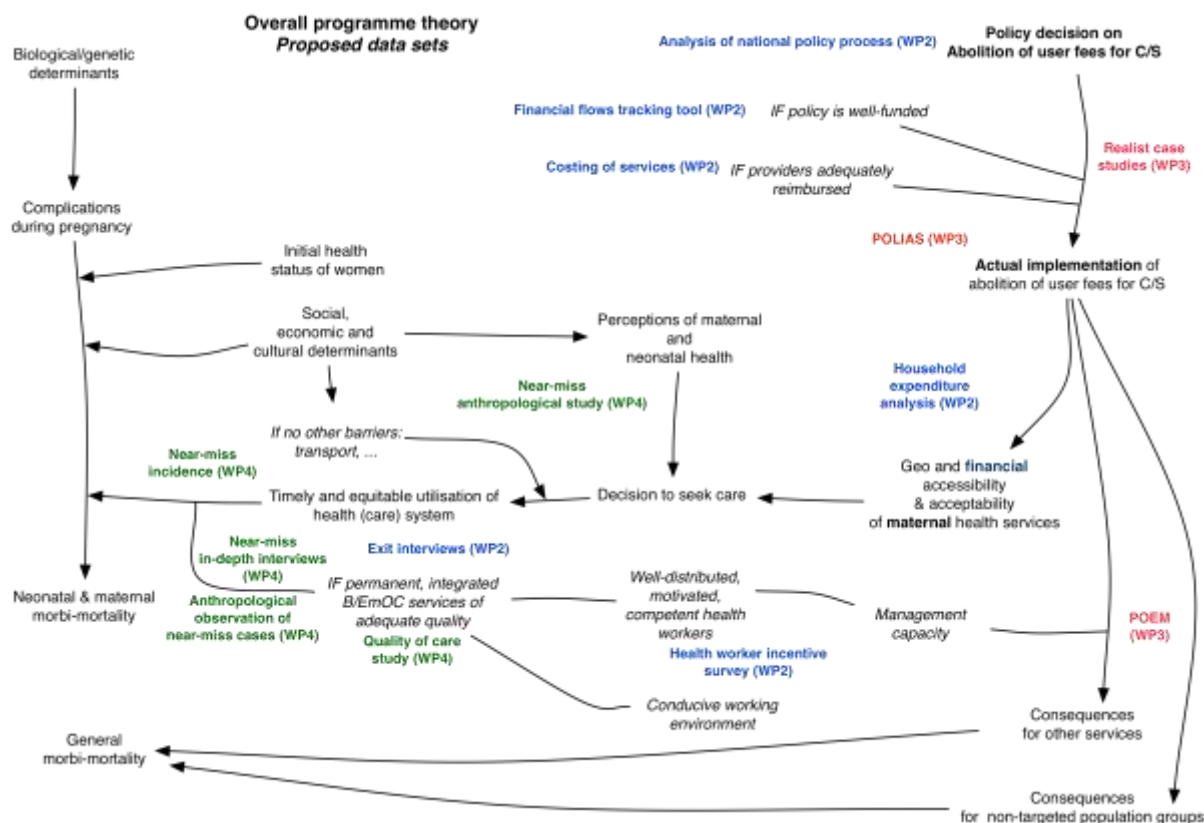


It presents the expected causal pathways at each level in the form of input-process-output-outcome configurations, but does not specifically address neither the linkages between the different levels nor the influence of the context. This is where the programme theory idea comes in. Although the FEMHealth programme did not set out as a theory-driven research project, we found it useful to develop a programme theory early on.

The concept of programme theory is central to theory-driven evaluation (see above). Chen and Rossi (1989) argue that for any intervention, a programme theory can be described that explains

how the planners expect the intervention to reach its objective. Describing the often implicit set of assumptions that steer the choice and design of an intervention allows us to understand what is being implemented and why. It should be noted that ‘theory’ is defined by Chen & Rossi as the “*prosaic theories that are concerned with how human organizations work and how social problems are generated*” (Chen and Rossi, 1983). The same can be done for a fee exemption policy. Figure 3 shows a simple version of the programme theory onto which we mapped the various sub-studies of FEMHealth. More detailed programme theories were developed to describe the effects of the policy on the local health system, to analyse the adoption and implementation of the policy by local service managers and providers, and to map how fee exemption would influence the health seeking pathways of pregnant women.

Figure 3 – Mapping the FEMHealth sub-studies⁷ onto a simple version of the overall programme theory



Developing a programme theory at the start of the programme served two goals. First, we aimed to facilitate a structured discussion among the researchers of their own hypotheses. As shown in Figure 2, these were developed for each level, but we felt that the work package format, favoured by the EC, posed a major risk of fragmentation. Large-scale research programmes such as FEMHealth are indeed typically organised in work packages, run by small teams of researchers, who tend to focus first on their specific research questions and only later (if time is available) on the overall objectives of the programme. Discussing the programme theory would lead, it was hoped, to better integration of the sub-studies.

⁷ The different colours represent different work packages within the research programme – WP2 focused on health policy and financing issues, WP3 on local health system issues, and WP4 on quality of care and utilisation responses.

In practice, the programme theory provided a framework to map the initially planned sub-studies, to find the blind spots and to better integrate the data collection. For instance, it allowed us to manage the gaps and overlaps between the work package focusing on the policy development process at national level and the team working on policy implementation within the districts. For both teams, the interface between policy/programme and implementation was important, and thinking through the transitions – from policy formulation to programme design and finally its implementation – helped us to be more efficient in data collection and in planning the analysis of data. In other cases, ideas for additional qualitative work emerged – for example, the relative absence of community-level research activities became apparent to the team - and the programme theory helped to frame this in the overall research programme.

The programme theory also proved useful to review the programme's progress and map and integrate emergent changes into the overall design. In short, it provided a common framework for seven research teams from six countries to collaborate on one overall research question. Finally, the overall programme theory was intended to provide a broad framework for cross-national comparison between the study countries. It would do so by drawing attention not only to the assessment of the actual implemented policy, but also to the specific contexts in which it took place and to the causal chains that linked the observed outcomes to the implemented policy. We discuss the main challenges we faced in the next section.

Studying a complex problem - the challenges and some possible solutions

While using concepts from complexity theory proved useful in broadening our view on fee exemption policies and in informing the overall research design, we encountered a number of challenges, only some of which were mitigated by using a programme theory perspective.

First, while the programme theory development proved useful in FEMHealth, it arguably came too late in the process. One of the strengths of using a theory-driven approach is that it demands a multi-disciplinary analysis. However, research funding mechanisms that fund such joint preparation processes during project design are rare. Indeed, most operate with tight deadlines that often preclude meetings and discussions among researchers on issues other than general outlines of a proposal. Second, most research proposal formats necessitate committing to a design and set of tools and 'deliverables' from the very proposal development phase. This meant that in our case, the process of developing the conceptual framework and the work on the overall programme theory followed, not preceded, the specification of research tools in the proposal submitted to the funder. We found that the approved protocol allowed reasonable margins of freedom in the sense that the programme outline, the deliverables and the time table were fixed, but that the demanded level of description of the work packages left sufficient leeway to adapt the protocol to new insights and results of preliminary data analysis.

Secondly, interdisciplinary teams seem natural to research and evaluation of complex interventions, but they demand particular attention to communication and debate. Coming from the disciplines of health economics, anthropology, midwifery, statistics, demography, public health and epidemiology, the FEMHealth researchers held quite diverse sets of assumptions on how to address the policy question. Building and refining the programme theory helped to make our assumptions clear and to better take them into account in the data collection and analysis phase. However, building in enough face-to-face engagement for all team members to be comfortable with the programme theory was a challenge. In addition to disciplinary differences, we also faced the challenge of working across two languages (English and French), and being distributed across different countries.

A third challenge facing researchers working on a complex problem is the sheer volume of information that is generated if all aspects of the issue need to be covered, and thus the capacity needed to collect and process this information. A typical (human) response is to reduce complexity and to artificially limit the scope to a feasible level. This is typically done in big research projects by

cutting up the issue into bits that are manageable by small research groups. During the proposal development phase of the FEMHealth project, each work package proposed a study design, with assorted methods and tools for data collection. The process of discussing the programme theory helped in reframing the protocols and data collection processes of these groups in the overall picture. As we are in the analysis phase at the time of writing, the programme theory has still to prove its usefulness in allowing integration of evidence from very diverse sub-studies. What is already clear, however, is the significant communication cost and the time required to bring together all relevant data and insights. The organisational capacity, limited project timeframe and competing demands on the time of the researchers are often major barriers to such integration. On a very practical level, since integration can by definition only happen at the end of the basic analysis phase, the fact that project results tend to be produced up against the deadlines often means that this (most rich) part of the analytical process is lost or unduly squeezed, in terms of time and attention.

A fourth challenge is to capture the significant relations and processes that lead to emergent behaviour in complex systems, or in the case of a policy implementation, the responses that result from the interaction between key actors and institutions in terms of structure and culture. To the extent that some of these responses are emergent and thus not predictable, total planning for data collection plan is not possible, and flexibility needs to be built in. In FEMHealth, the qualitative data collection process proved most useful to explore such emergent issues as these tools by definition maintain flexibility. These included, for example, policy ethnographies and interviews with key actors to document the interactions at the global-national interface and to describe the national-level policymaking processes. The policy implementation process was documented by a combination of methods, including interviews of actors at different levels of the health system. Another approach was to try to document outcome patterns and to explore the unexpected results through mixed research methods. To this end, we set out to assess the effects of the policy by a combination of measuring changing near-miss incidence, conducting observations in facilities, exit interviews and in-depth interviews with patients.

Yet another strategy to capture emergence was to use realist evaluation as the approach to the study of the policy adoption by health service managers and providers. A programme theory was developed on the basis of a literature review and tested in two sites in three of the countries. This sub-study reduced the complexity challenge by zooming in on a specific aspect of the policy implementation process, while at the same time allowing for a complexity perspective in the analysis of that aspect.

If flexible designs are required, the question of replicability arises. Whereas replicability of other kinds of studies relies mainly on the quality of the study protocol and the adherence to it, studies of complex issues that have an important emerging part need to ensure traceability of changes made to the protocol and to clearly document why changes were made in the first place.

In studies dealing with complex issues, the researchers need to adopt an adaptive attitude during the data collection to keep in tune with the evolving understanding of the issue under study. They need to be able to identify and capture unforeseen events that may be critical for the study. In other words, analytical capacity and research experience matters, as data collection through closed questionnaires and quantitative surveys will not be sufficiently flexible.

The programme theory may also prove helpful in this respect. Yin (1999) advocates the use of multiple cases in a replication process to enhance the theory-building capacity of this design: "*The remedy is to consider a case study, as a unit, to be equivalent to an experiment, as a unit; multiple-case studies may then be considered equivalent to multiple experiments.*" Yin argues that replication logic can be based on the theory behind the cases. In order to test this hypothesis, 'critical' cases are selected and their results compared on the basis of the initial hypothesis. If the same results are found and rival hypotheses can be eliminated, the theory is strengthened. Through this process of analytical generalisation, findings of case studies can thus be generalised

to the theoretical propositions (not to populations, as quasi-experimental methods attempt to do) (Yin, 2009). This is in line with the principles of theory-based evaluation (Chen, 1990, Chen, 2005).

In the case of FEMHealth, the challenge was to go beyond sub-study specific analysis and cut across databases to do a comprehensive analysis at country-level, and then at cross-national level. One solution is to dissolve work package groups and regroup researchers in country-specific teams that focus on integrative analysis of the cases. This is the stage in which the project finds itself at the writing of this paper. In the case of some research groups, researchers are releasing 'ownership' and reducing territorial behaviour over specific tools and datasets and start to break-down barriers and share knowledge. Allowing time for this to happen fully is now the challenge. A common solution is for researchers to continue to invest after project funding has ceased. This, however, demands a degree of institutional support and capacity, which is most challenging in both under-funded or overstretched research institutions in many low-income countries. It should be noted that the incentive structure in many if not all academic institutions, in the north and south, remains firmly geared towards individual production and profiling, and this constitutes a high barrier to collaboration.

Conclusion

Thinking in terms of 'simple-complicated-complex' and trying to understand where a problem or situation can be located in the sense-making frames helps in choosing the most appropriate designs and research methods. Such frames also point out where and when 'hard' evidence can be obtained, and subsequently used in decision-making, and when we need to accept that evidence as construed by evidence-based medicine and policy adepts can simply not be produced. In such cases, one can attempt to break down the complex issue into components, whereby we may need to accept that something will be lost in the simplification process. If that is not possible or desirable, adaptive approaches to research and evaluation are needed.

Perhaps the hardest thing is to see how in truly complex situations we can make sense of things, and to accept that complex issues can only be understood in retrospect. Only plausible explanations can be developed, not predictive theories. This does not mean that health policy and systems research reverts to journalism. Rather, building upon what is known and learning while doing becomes important. It is here that theory-driven inquiry approaches are most useful.

References

- ANDERSON, R. 2008. New MRC guidance on evaluating complex interventions. *BMJ*, 337, a1937.
- ARGYRIS, C. & SCHÖN, D. 1978. *Organization learning: A theory of action perspective* Reading, Mass, Addison Wesley.
- BEGUN, J. W., ZIMMERMAN, B. & DOOLEY, K. 2003. *Health care organisations as complex adaptive systems. Advances in health care organization theory. First ed.* San Francisco: Jossey-bass.
- BEN AMEUR, A., RIDDE, V., BADO, A. R., INGABIRE, M. G. & QUEUILLE, L. 2012. User fee exemptions and excessive household spending for normal delivery in Burkina Faso: the need for careful implementation. *BMC Health Serv Res*, 12, 412.
- BERMAN, P. 1978. *The study of macro and micro implementation of social policy*, Santa Monica, CA, RAND Corporation.
- BLAISE, P. & KEGELS, G. 2004. A realistic approach to the evaluation of the quality management movement in health care systems: a comparison between European and African contexts based on Mintzberg's organizational models. *Int J Health Plann Manage*, 19, 337-64.
- BOULTON, J. 2010. Complexity theory and implications for policy development. *Emergence: Complexity and Organisation*, 12, 31-40.
- BYNG, R., NORMAN, I. & REDFERN, S. 2005. Using realistic evaluation to evaluate a practice-level intervention to improve primary healthcare for patients with long-term mental illness. *Evaluation*, 11, 69-93.
- BYNG, R., NORMAN, I., REDFERN, S. & JONES, R. 2008. Exposing the key functions of a complex intervention for shared care in mental health: case study of a process evaluation. *BMC Health Serv Res*, 8, 274.
- BYRNE, D. 1998. *Complexity theory and the social sciences*, London, Routledge.
- CARASSO, B. S., LAGARDE, M., CHEELO, C., CHANSA, C. & PALMER, N. 2012. Health worker perspectives on user fee removal in Zambia. *Hum Resour Health*, 10, 40.
- CHANDRASEKARAN, P., DALLABETTA, G., LOO, V., MILLS, S., SAIDEL, T., ADHIKARY, R., ALARY, M., LOWNDES, C. M., BOILY, M. C. & MOORE, J. 2008. Evaluation design for large-scale HIV prevention programmes: the case of Avahan, the India AIDS initiative. *AIDS*, 22 Suppl 5, S1-15.
- CHEN, H. & ROSSI, P. 1983. Evaluating with sense - The theory driven approach. *Evaluation Review*, 7, 283-302.
- CHEN, H.-T. 1989. The conceptual framework of the theory-driven perspective. *Evaluation and Program Planning*, 12, 391-396.
- CHEN, H.-T. 1990. *Theory-driven evaluations*, Newbury Park, California, Sage Publications.
- CHEN, H.-T. 2005. *Practical program evaluation*, Thousand Oaks, SAGE Publications.
- CILLIERS, P. 1998. *Complexity and postmodernism. Understanding complex systems*, London and New York, Routledge.
- CONNELL, J. P., KUBISCH, A., SCHORR, L. & WEISS, C. 1995. *New approaches to evaluating community initiatives. Concepts, methods, and contexts*, Washington D.C., The Aspen Institute.
- CONNELL, J. P. & KUBITSCH, A. C. 1998. *Applying a theory of change approach to the evaluation of comprehensive community initiatives: progress, prospects, and problems. New approaches to evaluating community initiatives, Vol. 2.* Washington: Aspen Institute.
- CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I. & PETTICREW, M. 2008a. *Developing and evaluating complex interventions: new guidance.* Medical Research Council, UK.
- CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I. & PETTICREW, M. 2008b. *Developing and evaluating complex interventions: the new Medical Research Council guidance.* *BMJ*, 337, a1655.

- DE ALLEGRI, M., RIDDE, V., LOUIS, V. R., SARKER, M., TIENDREBEOGO, J., YE, M., MULLER, O. & JAHN, A. 2011. Determinants of utilisation of maternal care services after the reduction of user fees: a case study from rural Burkina Faso. *Health Policy*, 99, 210-8.
- DE BROUWERE, V., RICHARD, F. & WITTER, S. 2010. Access to maternal and perinatal health services: lessons from successful and less successful examples of improving access to safe delivery and care of the newborn. *Tropical Medicine & International Health*, 15, 901-9.
- DE SAVIGNY, D. & ADAM, T. 2009. *Systems thinking for health systems strengthening*. Geneva: Alliance for Health Policy and Systems Research & World Health Organisation.
- DENNARD, L., RICHARDSON, K. & MORCÖL, G. 2008. *Complexity and policy analysis. tools and methods for designing robust policies in a complex world*, Goodyear, AZ, ISCE Publishing.
- DENSCOMBE, M. 2003. *The good research guide for small-scale social research projects*, Maidenhead, Open University Press.
- DHILLON, R. S., BONDS, M. H., FRADEN, M., NDAHIRO, D. & RUXIN, J. 2012. The impact of reducing financial barriers on utilisation of a primary health care facility in Rwanda. *Glob Public Health*, 7, 71-86.
- DIELEMAN, M., GERRETSEN, B. & VAN DER WILT, G. J. 2009. Human resource management interventions to improve health workers' performance in low and middle income countries: a realist review. *Health Res Policy Syst*, 7, 7.
- DIXON-WOODS, M., AGARWAL, S., YOUND, B., JONES, D. K. & SUTTON, A. 2004. *Integrative approaches to qualitative and quantitative evidence*. London: NHS Health Development Agency.
- EL-KHOURY, M., HATT, L. & GANDAHO, T. 2012. User fee exemptions and equity in access to caesarean sections: an analysis of patient survey data in Mali. *Int J Equity Health*, 11, 49.
- FORSS, K., MARRA, M. & SCHWARTZ, R. 2011. *Evaluating the complex*, New Brunswick and London, Transaction Publishers.
- GARNETT, G. P., COUSENS, S., HALLETT, T. B., STEKETEE, R. & WALKER, N. 2011. Mathematical models in the evaluation of health programmes. *Lancet*, 378, 515-25.
- GARVIN, D. A. 1993. Building a learning organization. *Harv Bus Rev*, 71, 78-91.
- GILSON, L. 2012. *Health Policy and Systems Research. A methodology reader*, Geneva, Alliance for Health Policy and Systems Research & World Health Organisation.
- GILSON, L. & MCINTYRE, D. 2005. Removing user fees for primary care in Africa: the need for careful action. *BMJ*, 331, 762-5.
- GLOUBERMAN, S. & MINTZBERG, H. 2001. Managing the care of health and the cure of disease--Part I: Differentiation. *Health Care Manage Rev*, 26, 56-69; discussion 87-9.
- GLOUBERMAN, S. & ZIMMERMAN, B. 2002. *Complicated and complex systems: what would successful reform of Medicare look like? : Commission on the Future of Health Care in Canada*.
- GOICOLEA, I., COE, A. B., HURTIG, A. K. & SAN SEBASTIAN, M. 2012. Mechanisms for achieving adolescent-friendly services in Ecuador: a realist evaluation approach. *Global health action*, 5.
- GREENHALGH, T., HUMPHREY, C., HUGHES, J., MACFARLANE, F., BUTLER, C. & PAWSON, R. 2009. How do you modernize a health service? A realist evaluation of whole-scale transformation in London. *Milbank Q*, 87, 391-416.
- HERCOT, D., MEESEN, B., RIDDE, V. & GILSON, L. 2011. Removing user fees for health services in low-income countries: a multi-country review framework for assessing the process of policy change. *Health Policy Plan*, 26 Suppl 2, ii5-15.
- HM TREASURY 2011. *The Magenta Book. Guidance for evaluation*. London: HM Treasury, UK.
- HOMER, J. B. & HIRSCH, G. B. 2006. System dynamics modeling for public health: background and opportunities. *American Journal of Public Health*, 96, 452-8.
- IDD, A., YOHANA, O. & MALUKA, S. O. 2013. Implementation of pro-poor exemption policy in Tanzania: policy versus reality. *Int J Health Plann Manage*, Epub ahead of print.

- JONES, H. 2011. Taking responsibility for complexity. How implementation can achieve results in the face of complexity. Working Paper. London: Overseas Development Institute.
- JUDGE, K. & BAULD, L. 2006. Learning from policy failure? Health action zones in England. *Eur J Public Health*, 16, 341-3.
- KANE, S., GERRETSEN, B., SCHERPBIER, R., DAL POZ, M. & DIELEMAN, M. 2010. A realist synthesis of randomised control trials involving the use of community health workers for delivering child health interventions in low and middle income countries. *BMC Health Services Research*, 10, 286.
- KERNICK, D. 2002. The demise of linearity in managing health services: a call for post normal health care. *J Health Serv Res Policy*, 7, 121-4.
- KERNICK, D. 2004. Complexity and healthcare organisation, Oxford, Radcliffe Medical Press.
- KERNICK, D. 2006. Wanted--new methodologies for health service research. Is complexity theory the answer? *Fam Pract*, 23, 385-90.
- KINGDON, J. 1995. Agendas, alternatives and public policies, Addison-Wesley Educational Publishers Inc.
- KOLB, D. 1984. Experiential learning: experience as the source of learning and development New Jersey, Prentice-Hall Englewood Cliffs.
- KURTZ, C. & SNOWDEN, D. 2003. The new dynamics of strategy: sense-making in a complex and complicated world. *IBM Systems Journal*, 42, 462-483.
- LAGARDE, M. 2012. How to do (or not to do) ... Assessing the impact of a policy change with routine longitudinal data. *Health Policy Plan*, 27, 76-83.
- LAGARDE, M., BARROY, H. & PALMER, N. 2012. Assessing the effects of removing user fees in Zambia and Niger. *J Health Serv Res Policy*, 17, 30-6.
- LEISCHOW, S. J. & MILSTEIN, B. 2006. Systems thinking and modeling for public health practice. *American Journal of Public Health*, 96, 403-5.
- LEONE, L. 2008. Realistic evaluation of an illicit drug deterrence programme. *Evaluation*, 14, 19-28.
- LINDBLOM, C. 1980. The policymaking process, New Jersey, Prentice-Hall.
- LINSTEAD, S., FULOP, L. & LILLEY, S. 2009. Management & Organization. A critical text, London, Palgrave MacMillan.
- MACKENZIE, M., KOSHY, P., LESLIE, W., LEAN, M. & HANKEY, C. 2009. Getting beyond outcomes: a realist approach to help understand the impact of a nutritional intervention during smoke cessation. *European Journal of Clinical Nutrition*, 63, 1136-1142.
- MALUKA, S., KAMUZORA, P., SAN SEBASTIAN, M., BYSKOV, J., NDAWI, B., OLSEN, O. & HURTIG, A. K. 2011. Implementing accountability for reasonableness framework at district level in Tanzania: a realist evaluation. *Implementation Science*, 6.
- MANZANO-SANTAELLA, A. 2011. A realistic evaluation of fines for hospital discharges: Incorporating the history of programme evaluations in the analysis. *Evaluation*, 17, 21-36.
- MARCH, J. & OLSEN, J. 1976. Ambiguity and choice in organisations, Oslo, Universitetsforlaget.
- MARCHAL, B., DEDZO, M. & KEGELS, G. 2010a. A realist evaluation of the management of a well-performing regional hospital in Ghana *BMC Health Services Research*, 10.
- MARCHAL, B., DEDZO, M. & KEGELS, G. 2010b. Turning around an ailing district hospital: a realist evaluation of strategic changes at Ho Municipal Hospital (Ghana). *BMC Public Health*, 10, 787.
- MARCHAL, B., VAN BELLE, S., DE BROUWERE, V. & WITTER, S. 2013. Studying complex interventions: reflections from the FEMHealth project on evaluating fee exemption policies in West Africa and Morocco. *BMC Health Services Research*.
- MARCHAL, B., VAN BELLE, S., VAN OLMEN, J., HOERÉE, T. & KEGELS, G. 2012. Is realist evaluation keeping its promise? A literature review of methodological practice in health systems research. *Evaluation*, 18, 192-212.

- MAYNE, J. 2001. Addressing attribution through contribution analysis: using performance measures sensibly. *Canadian Journal of Program Evaluation*, 16, 1-24.
- MAYNE, J. 2011. Contribution analysis: addressing cause and effect. In: FORSS, K., MARRA, M. & SCHWARTZ, R. (eds.) *Evaluating the complex: attribution, contribution, and beyond*. New Brunswick: Transaction Publishers.
- MAYNE, J. 2012. Contribution analysis: Coming of age? *Evaluation*, 18, 270-280.
- MCDANIEL, R., LANHAM, H. & ANDERSON, R. 2009. Implications of complex adaptive systems theory for the design of research on health care organizations. *Health Care Management Review*, 34, 191-199.
- MCPAKE, B., BRIKCI, N., COMETTO, G., SCHMIDT, A. & ARAUJO, E. 2011. Removing user fees: learning from international experience to support the process. *Health Policy Plan*, 26 Suppl 2, ii104-117.
- MEEK, J. 2010. Complexity theory for public administration and policy. *Emergence: Complexity and Organisation*, 12, 1-4.
- MEESSEN, B., HERCOT, D., NOIRHOMME, M., RIDDE, V., TIBOUTI, A., TASHOBYA, C. K. & GILSON, L. 2011. Removing user fees in the health sector: a review of policy processes in six sub-Saharan African countries. *Health Policy Plan*, 26 Suppl 2, ii16-29.
- MIDGLEY, G. 2006. Systemic intervention for public health. *American Journal of Public Health*, 96, 466-72.
- MILTON, B., MOONAN, M., TAYLOR-ROBINSON, D. & WHITEHEAD, M. 2011. How can the health equity impact of universal policies be evaluated? Insights into approaches and next steps. Copenhagen: World Health Organisation - Regional Office for Europe
- MINISTÈRE DE LA SANTÉ 2006. *Stratégie nationale de subvention des accouchements et des soins obstétricaux et neonatals d'urgence au Burkina Faso*. Ougadougou: Direction de la Santé de la Famille.
- MORÇÖL, G. 2010. Issues in reconceptualizing public policy from the perspective of complexity theory. *Emergence: Complexity and Organisation*, 12, 52-60.
- MORIN, E. 2001. *Seven complex lessons in education for the future*. Paris: UNESCO Publishing.
- NABYONGA OREM, J., KAAWA MAFIGIRI, D., MARCHAL, B., SSENGOOBA, F., MACQ, J. & CRIEL, B. 2012. Research, evidence and policymaking: the perspectives of policy actors on improving uptake of evidence in health policy development and implementation in Uganda. *BMC Public Health*, 12, 109.
- NABYONGA OREM, J., MARCHAL, B., D., M., SSENGOOBA, F., MACQ, J., CAMPOS DA SILVEIRA, V. & CRIEL, B. 2013. Perspectives on the role of stakeholders in knowledge translation in health policy development in Uganda. *BMC Health Services Research*, 13, 324.
- NG, M., GAKIDOU, E., LEVIN-RECTOR, A., KHERA, A., MURRAY, C. J. & DANDONA, L. 2011. Assessment of population-level effect of Avahan, an HIV-prevention initiative in India. *Lancet*, 378, 1643-52.
- NIMPAGARITSE, M. & BERTONE, M. P. 2011. The sudden removal of user fees: the perspective of a frontline manager in Burundi. *Health Policy Plan*, 26 Suppl 2, ii63-71.
- OSTROM, E. 1990. *Governing the commons: the evolution of institutions for collective action*, Cambridge, Cambridge University press.
- PATTON, M. Q. 2011. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, New York.
- PAWSON, R. 2002. Evidence-based policy: the promise of 'realist synthesis'. *Evaluation*, 8, 340-358.
- PAWSON, R. 2006. *Evidence-based policy. A realist perspective*, London, Sage.
- PAWSON, R., GREENHALGH, T., HARVEY, G. & WALSHE, K. 2005. Realist review--a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy*, 10 Suppl 1, 21-34.
- PAWSON, R. & TILLEY, N. 1997. *Realistic Evaluation*, London, Sage.
- PERKINS, M. B., JENSEN, P. S., JACCARD, J., GOLLWITZER, P., OETTINGEN, G., PAPPADOPULOS, E. & HOAGWOOD, K. E. 2007. Applying theory-driven approaches to understanding and modifying clinicians' behavior: what do we know? *Psychiatr Serv*, 58, 342-8.

- PETTIGREW, A. 1990. Longitudinal field research on change: theory and practice. *Organization science*, 1.
- PETTIGREW, A. 1997. What is processual analysis? *Scandinavian Journal of Management*, 13, 337-348.
- PLSEK, P. 2001. Redesigning health care with insights from the science of complex adaptive systems. *Crossing the quality chasm: a new health system for the 21st century*. Washington, D.C.: Committee on Quality of Health Care in America, Institute of Medicine, The National Academy of Sciences.
- PLSEK, P. & GREENHALGH, T. 2001. The challenge of complexity in health care. *British Medical Journal*, 323, 625-628.
- PRASHANTH, N., MARCHAL, B., HOERÉE, T., DEVADASAN, N., MACQ, J., KEGELS, G. & CRIEL, B. 2012. Building capacity of health managers: How does it work? A realist evaluation study. *BMJ Open*, 2, e000882.
- PRIESMEYER, H. R., SHARP, L. F., WAMMACK, L. & MABREY, J. D. 1996. Chaos theory and clinical pathways: a practical application. *Quality management in health care*, 4, 63-72.
- RAMALINGAM, B. & JONES, H. 2008. Exploring the science of complexity: Ideas and implications for development and humanitarian efforts. Working Paper. London: Overseas Development Institute.
- REED, M. & HARVEY, D. 1992. The new science and the old: complexity and realism in the social sciences. *Journal for the Theory of Social Behaviour*, 22, 356-79.
- RESNICOW, K. & PAGE, S. E. 2008. Embracing chaos and complexity: a quantum change for public health. *Am J Public Health*, 98, 1382-9.
- RICHARD, F., ANTONY, M., WITTER, S., KELLEY, A., MEESSEN, B. & SIELEUNOU, I. 2013. Fee exemption for maternal care in Sub-Saharan Africa: a review of 11 countries and lessons for the region. *Global Health Governance*, 6,, 1-21.
- RICHARD, F., WITTER, S. & DE BROUWERE, V. 2010. Innovative approaches to reducing financial barriers to obstetric care in low-income countries. *American Journal of Public Health*, 100, 1845-52.
- RICKLES, D. 2008. Causality in complex interventions. *Med Health Care Philos*, 12, 77-90.
- RICKLES, D., HAWE, P. & SHIELL, A. 2007. A simple guide to chaos and complexity. *J Epidemiol Community Health*, 61, 933-7.
- RIDDE, V. 2011. What do we know in 2011 about the new free care policy in Mali? Montréal, Canada.
- RIDDE, V. & DIARRA, A. 2009. A process evaluation of user fees abolition for pregnant women and children under five years in two districts in Niger (West Africa). *BMC Health Services Research*, 9, 89.
- RIDDE, V., HADDAD, S., NIKIEMA, B., OUEDRAOGO, M., KAFANDO, Y. & BICABA, A. 2010. Low coverage but few inclusion errors in Burkina Faso: a community-based targeting approach to exempt the indigent from user fees. *BMC Public Health*, 10, 631.
- RIDDE, V. & MORESTIN, F. 2011. A scoping review of the literature on the abolition of user fees in health care services in Africa. *Health Policy and Planning*, 26, 1-11.
- ROBERT, E., RIDDE, V., FOURNIER, P. & BICABA, A. 2010. How to use the Realist approach to inform policy-making: the example of a research method on user fees abolition policies in Africa. 2010 symposium of Canadian Cochrane Centre.
- ROBERT, E., RIDDE, V., MARCHAL, B. & FOURNIER, P. 2012. Protocol: a realist review of user fee exemption policies for health services in Africa. *BMJ Open*, 2, e000706.
- ROGERS, P. 2009. Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation. In: WHITE, H. (ed.) *Designing impact evaluations: different perspectives*. International Initiative for Impact Evaluation Working Paper 4. New Delhi: 3ie.
- ROGERS, P. 2011. Implications of complicated and complex characteristics for key tasks in evaluation. In: FORSS, K., MARRA, M. & SCHWARTZ, R. (eds.) *Evaluating the complex. Attribution, contribution, and beyond*. New Brunswick, London: Transaction Publishers.
- SABATIER, P. 1999. *Theories of the policy process*, Boulder, Colorado, Westview Press.

- SABATIER, P. & JENKINS-SMITH, H. 1993. Policy change and learning: an advocacy coalition approach, Boulder, CO, Westview press.
- SABATIER, P. & MAZMANIAN, D. 1979. The implementation of public policy: a framework of analysis.
- SANDERSON, I. 2002. Evaluating complexity: theory, evidence and practical guidance.
- SCHÖN, D. 1983. The Reflective Practitioner, How Professionals Think In Action, New York, Basic Books.
- SCHWARTZ, R. & GARCIA, J. 2011. Intervention Path Contribution Analysis (IPCA) for complex strategy evaluation: evaluating the smoke-free Ontario strategy. In: FORSS, K., MARRA, M. & SCHWARTZ, R. (eds.) Evaluating the complex. Attribution, contribution, and beyond. New Brunswick (USA): Transaction Publishers.
- SENGE, P. 1990. The fifth discipline, New York, Currency Doubleday.
- SIDANI, S., DORAN, D. M. & MITCHELL, P. H. 2004. A theory-driven approach to evaluating quality of nursing care. *J Nurs Scholarsh*, 36, 60-5.
- SNOWDEN, D. & BOONE, M. 2007. A leader's framework for decision making. *Harvard Business Review*.
- STACEY, R. 1996a. Complexity and creativity in organisations, San Francisco, Berret-Koehler.
- STACEY, R. 1996b. Strategic management and organizational dynamics, London, Pitman Publishing.
- STERMAN, J. D. 2006. Learning from evidence in a complex world. *Am J Public Health*, 96, 505-14.
- SWANSON, D. & BHADWAL, S. 2009. Creating adaptive policies - A guide for policymaking in a uncertain world, New Delhi, SAGE Publications.
- TOLSON, D., MCINTOSH, J., LOFTUS, L. & CORMIE, P. 2007. Developing a managed clinical network in palliative care: a realistic evaluation. *Int J Nurs Stud*, 44, 183-95.
- TROCHIM, W. M., CABRERA, D. A., MILSTEIN, B., GALLAGHER, R. S. & LEISCHOW, S. J. 2006. Practical challenges of systems thinking and modeling in public health. *American Journal of Public Health*, 96, 538-46.
- VAN BELLE, S., MARCHAL, B., DUBOURG, D. & KEGELS, G. 2010. How to develop a theory-driven evaluation design - Lessons learned from an adolescent sexual and reproductive health programme in West-Africa. *BMC Public Health*, 10, 741.
- VICTORA, C. G., BLACK, R. E., BOERMA, J. T. & BRYCE, J. 2011. Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations. *Lancet*, 377, 85-95.
- WALKER, L. & GILSON, L. 2004. 'We are bitter but we are satisfied': nurses as street-level bureaucrats in South Africa. *Soc Sci Med*, 59, 1251-61.
- WESTHORP, G. 2012. Using complexity-consistent theory for evaluating complex systems. *Evaluation*, 18, 405-420.
- WHITE, H. & PHILLIPS, D. 2012. Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework. New Delhi: International Initiative for Impact Evaluation.
- WILSON, T. & HOLT, T. 2001. Complexity and clinical care. *British Medical Journal*, 323, 685-688.
- WITTER, S. 2009. Service- and population-based exemptions: are these the way forward for equity and efficiency in health financing in low income countries? . In: CHERNICHOVSKY, D. & HANSON, K. (eds.) *Advances in Health Economics and Health Services Research*. Innovations in health systems finance in developing and transitional economies. Emerald.
- WITTER, S. & ADJEI, S. 2007. Start-stop funding, its causes and consequences: a case study of the delivery exemptions policy in Ghana. *The International Journal of Health Planning and Management*, 22, 133-43.
- WITTER, S., ARHINFUL, D. K., KUSI, A. & ZAKARIAH-AKOTO, S. 2007a. The experience of Ghana in implementing a user fee exemption policy to provide free delivery care. *Reproductive Health Matters*, 15, 61-71.

- WITTER, S., DIENG, T., MBENGUE, D., MOREIRA, I. & DE BROUWERE, V. 2010. The national free delivery and caesarean policy in Senegal: evaluating process and outcomes. *Health Policy and Planning*, 25, 384-92.
- WITTER, S., GARSHONG, B. & RIDDE, V. 2013. An exploratory study of the policy process and early implementation of free NHIS coverage for pregnant women in Ghana *International Journal for Equity in Health* 12, 1-11.
- WITTER, S., KHADKA, S., NATH, H. & TIWARI, S. 2011. The national free delivery policy in Nepal: early evidence of its impact on health facilities. . *Health Policy and Planning*, 26, 84-91.
- WITTER, S., KHALID, K., ABDEL-RAHMAN, M., HUSSEIN, R. & SAED, M. 2012. Removal of user fees for caesareans and under-fives in northern Sudan: a review of policy implementation and effectiveness. *International Journal of Health Planning and Management*, 28, e95-e120.
- WITTER, S., KUSI, A. & AIKINS, M. 2007b. Effects of a delivery exemption scheme on workloads and incomes of health workers in Ghana. *Human Resources for Health* 5.
- WITTER, S., RICHARD, F. & DE BROUWERE, V. 2008. Learning lessons and moving forward: how to reduce financial barriers to obstetric care in low-income countries. *Reducing financial barriers to obstetric care in low-income countries*. Antwerp: ITGPress.
- WONG, G., GREENHALGH, T. & PAWSON, R. 2010. Internet-based medical education: a realist review of what works, for whom and in what circumstances. *BMC Med Educ*, 10, 12.
- YIN, R. 1999. Enhancing the quality of case studies in health services research. *Health Services Research*, 34, 1209-1224.
- YIN, R. 2003. *Case study research. Design and methods*, London, Sage Publications.
- YIN, R. 2009. *Case study research. Design and methods*, Los Angeles, Sage.
- ZIMMERMAN, B., DUBOIS, N., HOULE, J., LLOYD, S., MERCIER, C., BROUSSELLE, A. & REY, L. 2012. How does complexity impact evaluation? An introduction to the special issue. *The Canadian Journal of Program Evaluation*, 26, v-xx.
- ZIMMERMAN, B., LINDBERG, C. & PLSEK, P. (undated) *A complexity primer: what is complexity science and why should I learn about it?*
- ZIMMERMAN, B., LINDBERG, C. & PLSEK, P. 1998. *Edgware: Lessons from complexity science for health care leaders*, Irving, Texas, Plexus Institute.