Addressing equity issues in monitoring and evaluation

Greet Peersman and Patricia Rogers, Better Evaluation, RMIT University

Introduction

The 'worst-off' have been targeted in international development interventions for decades, including the current push towards evaluating progress in achieving the Sustainable Development Goals (SDGs) using an equity-focused and gender-responsive lens (see, for example, Bamberger and Segone 2011; UNICEF 2011). Even in developed nations, the rhetoric of politicians, organisations and programs has often stated 'leave no one behind'. However, while we believe that there is genuine interest and willingness in achieving equitable development results, monitoring and evaluation (M&E) often fails to address equity issues in its findings and processes. In doing so, it fails to adequately inform efforts to improve equity and, worse, can undermine them.

While it is heartening to see equity being increasingly emphasised, most address equity in M&E in terms of approaches for 'increasing participation' and 'broadening measurement' (Rogers 2016). There is an assumption that 'scaling up' these approaches through strengthening national capacities will make a difference. We argue that this falls far short of what needs to be done.

There is more to M&E than choosing indicators and data collection methods; every choice about evaluation processes and methods has potential equity implications. We have found it helpful to think through each of the tasks involved in M&E as outlined in the BetterEvaluation Rainbow Framework (betterevaluation.org), from developing a clear statement about what is being evaluated and how it is understood to work (its theory of change), to framing the evaluation in terms of its primary focus and value framework, collecting and analysing data to answer descriptive, causal and evaluative questions, reporting findings and supporting their use, and all the processes involved in managing a single evaluation and an entire evaluation system including establishing standards for quality M&E and strengthening M&E capacity.

Working through the Rainbow Framework makes it clear that many equity issues need to be taken into consideration throughout the process of planning, managing, executing and using M&E (Rogers 2016). Those involved in M&E should be transparent about what was addressed or not and why, and, what the equity implications are for data collection, analysis and use. This chapter sets out what we see as the range of issues that need attention in order to address equity issues adequately and the implications for strengthening national M&E capacity.

Whose theory of change?

Many interventions are based implicitly or explicitly on a theory of change (ToC)—a statement about how the intervention activities are understood to contribute to a chain of results that ultimately lead to intended longer-term outcomes (or impact). Few, however, *directly* address equity issues in the ToC or in how it is developed and represented. Typically, outcomes refer to changes in the population—not prioritising improvements for the most disadvantaged—the change theory does not address how equity might be achieved, and the process for developing the ToC does not include those whose needs are supposedly being addressed (see, for example, http://www.csaf.net/impact/theory-of-change/; http://gadnetwork.org/gadn-resources/2015/3/6/untangling-gender-mainstreaming-a-theory-of-change-based-on-experience-and-reflection). The result is a ToC with a generic approach to bringing about change (e.g. increasing skills and knowledge of participants) which does not address the very real possibility that the less-disadvantaged will find it easier to access the program, staff will favour those whose needs are easier to meet, and overall improvements will mask that the most disadvantaged have fallen even further behind.

Framing M&E—users, uses and questions

M&E is about tracking progress of interventions towards intended goals, alerting to and understanding problems (including implementation challenges and unintended consequences) and judging the performance and merit of interventions. Hence, clarifying the values on which these judgements will be made is crucial. However, different stakeholders most often have different values and whose values prevail, sets a particular course of action for M&E that may be hard to change later on.

M&E is done for different reasons. Sometimes, merely to comply with a requirement, but usually it is for a particular purpose. Encouraging identification and prioritisation of primary intended users and uses is considered good practice (Quinn Patton 2008). However, the process itself risks increasing the powerlessness of the most disadvantaged when funders and/or managers of M&E do not consider the implications of privileging their own information needs in designing the M&E system. Whose learning and what types of accountability is M&E intended to support? What processes are in place to ensure that the most powerful do not have total control of the M&E resources to bolster their position?

Examples of evaluation questions that explicitly take equity issues into account are:

- How was the United States Institute of Peace Evaluation (USIP) theory of change developed with regards to minorities in Iraq and how did it shift over the course of the work? How are the Alliance of Iraqi Minorities and the Caucus different from how they were initially envisioned and what were the key decisions that led to these changes? (USIP 2013).

- To what extent have results contributed to decreased inequities between the best-off and the worst-off groups?

Answering descriptive questions—going beyond the average effect

Methods and processes for answering descriptive questions—factual statements about the way things are or were—must ensure that data from the most disadvantaged are included and analysis goes beyond looking at the average effect. This is a challenge when the disadvantaged are likely to be the hardest to reach and engage.

Disaggregating data to look for patterns across groups is not an optional 'nuance' but an essential element of doing ethical M&E, as the following example from Westhorp (2008) makes clear. A review of early intervention programs for children in disadvantaged families found that some programs which were effective on average, were either ineffective or damaging for some, notably those who often had multiple and complex needs or were amongst the most disadvantaged families. In particular, the Early Head Start program was found to have unfavourable impacts on child development outcomes in families with multiple risk factors (Westhorp 2008; Mathematica Policy Research Inc. 2002). The findings were reported in over-simplified ways and the reputable Brookings Institute, uncritically, recommended its wider adoption because it was 'evidence-based' (see: http://www.brookings.edu/~/media/Research/Files/Papers/2008/9/early-programs-isaacs/09_early_programs_brief2.PDF).

Answering causal questions—implications of privileging certain research designs

Causal questions such as why things are the way they are sometimes involve a counterfactual or an estimate of what would have happened in the absence of the intervention, but can also involve two other strategies: checking the data for consistency with the ToC and identifying and ruling out alternative explanations (see: http://betterevaluation.org/plan/understandcauses). However, continuing paradigm wars have seen a privileging of particular designs which has equity implications. In particular, some influential approaches to evidence-based policy, including the Coalition for Evidence-Based Policy (see: http://coalition4evidence.org/) and the Poverty Action Lab (https://www.povertyaction-lab.org/) have argued for the use of the randomised controlled trial (RCT) as 'the gold standard'—even to the point of only considering evidence as credible when studies have used this design.

Such studies require the design to be fixed at the start, and the groups to be disaggregated identified at the start and taken into account when calculating sample sizes. Where new, unexpected issues arise, the design cannot accommodate them, and it is not possible to identify patterns of results for new groups which were not evident at the beginning. Also, these designs can only be used for interventions that are designed for a unit of analysis that can be randomised. Hence, it privileges interventions at the level of an individual person, a classroom, a school or a village

which can be isolated from its control group 'neighbours' (see also Stern et al. 2012). By extension of the logic used by hardline RCT advocates, evaluations of interventions that cannot readily use a RCT design (e.g. interventions that are flexible and adaptive to changing needs or circumstances or that operate on the scale of a country, large region or service system) cannot provide evidence that can demonstrate their effectiveness.

Answering evaluative questions

Answering evaluative questions (whether something is good or bad, better or worse, good or poor value) involves identifying evaluative criteria and standards, gathering relevant data, and synthesising these data through weighing or trading off different criteria. Important equity considerations are:

- whose values are taken into account when determining evaluative criteria, standards and synthesis?
- who decides 'what success looks like'?
- should 'success' be considered an overall improvement or a reduction in equity gaps? Should it be equality of opportunity or of outcomes?
- if a program is good for some people and harmful for others, how should it be judged overall?

Reporting findings and supporting use

The final and perhaps, the most important, stage of M&E also has equity implications. There can be issues in terms of who receives reports about the findings, and how accessible these are. In some cases, new forms of reporting, such as video reports, theatrical performances and individual conversations, can make findings accessible to a wider group of community members (see, for example: http://better-evaluation.org/plan/reportandsupportuse/ report). The style of the report can also be important. Deficit-focused reporting can increase the stigma for a disadvantaged group or community (Gibney et al. 1999). One of the major challenges in understanding the effects of interventions, especially in the published literature, has been the lack of detail in reporting demographic and other relevant characteristics of the participants (Peersman and Rugg 2004; Welch et al. 2016). This impedes making informed decisions about continuing or improving the intervention or judging its potential applicability for replication elsewhere.

The increased availability of systematic reviews has provided important lessons that can be learned *across* different studies (Gough et al. 2012). However, the findings of an 'effectiveness review' have been shown to depend on how the review is done e.g. how studies were sought such as including of grey literature or studies reported in non-English journals; what criteria where used to include studies such as particular study designs (see, for example, Peersman et al. 1999). Important improvements to review methods have been made over the years, for example, the integration of qualitative data and increased 'consumer' involvement (Oliver and Peersman 2001; Gough et al. 2012). Most recently, a checklist of 'preferred reporting items' was published to promote a focus on health equity

(defined as 'avoidable and unfair differences in health outcomes') in systematic reviews (Welch et al. 2016). This provides a good starting point for increasing transparency and comprehensiveness in reporting differences in access, opportunity and outcomes between different groups.

Managing M&E—who makes the decisions on whose values count?

There are many decisions to be made in M&E. Control may be centralised in a dedicated manager, unit or committee or it may be shared by a working group or other entity involving representatives from many different stakeholders. It is important to describe clearly each actor's role to avoid confusion, duplication of effort or things falling through the cracks. Important equity questions are:

- Who makes decisions and what is their role—consulting, recommending, approving, is informed and how are decisions made—majority, consensus or hierarchical decision making?
- Does the planning of M&E activities involve values clarification and negotiation between groups informed by a multitude of sources (such as individuals' experiences and aspirations, research and comparative analysis)?
- Who controls the resources for M&E?
- Who decides on priority information needs or on what gets evaluated?
- Who collects and analyses the data? Who makes the evaluative judgments?
- What is the basis for judging the quality of an evaluation?

Table 1 provides a summary overview of the equity issues discussed above, using the Rainbow Framework of M&E Tasks.

Table 1. Equity issues across the range of M&E tasks

Cluster of M&E Tasks	Examples of equity considerations
Define—Developing or obtaining a description of the intervention and how it is understood to work. • What is the theory of change (ToC)? • What are possible unintended results?	° Who decides what the intended results are? ° Does the ToC just look at average results? ° Does the ToC have a change theory for improving equity? ° Which groups are specifically addressed in the ToC? Which equity issues (if any) are identified? ° Who decides what are the unintended, positive/negative, noticed/ unnoticed results?
Frame—Setting the parameters for M&E—the purposes, what to monitor and what to evaluate including key evaluation questions and information needs for decision-making about the intervention, and, the criteria and standards to be used. • What are the key questions M&E needs to address? • Who are the primary intended users of the M&E data?	° Do the M&E questions explicitly include equity considerations? ° Are staff, managers and policy-makers claiming to act on behalf of those who are disadvantaged? Or those directly experiencing inequity? ° Are the power structures traditional, collaborative or devolved? ° Is the primary purpose account-ability? Who is being held accountable for what and by whom?
• What are the primary intended uses of the M&E data?	° Is the primary purpose learning? Whose learning is being supported? Through what process? In terms of what functions?
Describe—Collecting or collating data to answer descriptive questions about the intervention, the various results observed, and the context in which the intervention is implemented. • What differences are described? How are they described?	For example: ° *Inputs*: limited access to services; barriers such as cost ° *Results*:different health or employment outcomes for people from particular ethnic groups
Understand causes—Collecting and analysing data to answer causal questions about the extent to which the intervention produced observed results. • What are the implications of the causal inference strategies used?	° Do the causal inference strategies allow for explanation of differential impacts? Do they skew findings to individual-level interventions?
Synthesize—Combining data to form an overall assessment of the merit, worth and/or significance of an intervention, or to summarise evidence across several evaluations. • How are evaluative criteria and standards decided (i.e., what does success looks like?)?	° Are interventions entirely structured around externally-set criteria and standards or the stated values of those experiencing inequity? ° Do M&E activities involve values clarification and negotiation between groups informed by a multitude of sources? ° Should the aim be equality of opportunity or equality of outcomes? What would be a reasonable trade-off between means and ends? ° If there are differential effects, how should these be synthesized? ° Should equity impacts become a 'hurdle' requirement when assessing overall success?
Report and support use—Developing and presenting findings in ways that are useful for the primary intended users, and supporting them to make evidence-informed decisions.	° Do deficit-focused reports increase stigma and deficit focus? ° How accessible are reports in terms of language, style, format, type etc.? ° What support is provided to whom to apply or use the findings?
Manage—The planning and management of the implementation of the M&E system, including who will make decisions about it, who will lead development and implementation and the roles and responsibilities of different actors.	° What gets evaluated? Welfare or subsidies? Short projects, ongoing programs or private sector products? ° How are the beneficiaries of the intervention engaged in decision making about M&E? Do they have both the opportunity and the capacity to be meaningfully and actively engaged? ° Who controls the resources for M&E? ° Who collects the data? Who analyses the data? Who makes the evaluative judgments? ° What is the basis for judging the quality of the evaluation?

Source: Adapted from: Peersman G, P Rogers. Addressing equity issues in monitoring and evaluation. 5th ACFID University Network Conference, Melbourne 4–5 June 2015.

Strengthening national M&E capacity to help reduce inequity

The issues already discussed are relevant for individual M&E projects or products. We now discuss the implications for strengthening national M&E systems, which are import-ant in evaluating the new SDGs. They are:

- the need for more thoughtful identification and use of equity categories;

- ways to support genuine engagement of the less powerful in M&E processes; and, what can be learned from previous efforts to strengthen national M&E capacities as part of evaluating the Millennium Development Goals (MDGs). We propose some minimum standards for what should be addressed in equity-focused M&E and what is required for effective M&E systems.

How 'equity' and 'gender' are defined matters.

Gender is a current 'buzz word' in international development as well as the focus of many governments' investments. Yet, gender-focused work tends to be based on apparently fixed male and female characteristics and assumed to be synonymous with 'working with women' (Fletcher 2015). Fletcher argues that '…to effectively address inequity and injustice, it [international development] needs to move away from categories and pay more attention to the shifting and intersecting processes of understanding and judgement related to gender, sexuality, and other intersecting social hierarchies such as class/caste or race/ethnicity' (Fletcher 2015: 1). Similarly, Bustelo et al. (2015:6) emphasises that activists and professionals need to pay attention to these inter-sectionalities and how they play out in terms of 'privileges and exclusions' to the detriment of some individuals or groups. Given the significant potential to do more harm than good, we should set minimum standards for interventions and their M&E to be called 'equity-focused' and 'gender-responsive'.

The Gender Equality Continuum makes a distinction between gender blind (absence of any pro-active consideration of the larger gender environment) and gender aware (deliberately examining and addressing gender in design and implementation). In the latter category, approaches can be classified as gender 'exploitative' (such as protection of the status quo), gender 'accommodating' (such as not challenging unequal relations of power) and gender 'transformative' (explicitly engaging women and men to examine, question and change institutions and norms that reinforce gender inequalities) (Interagency Gender Working Group 2009; see http://www.igwg.org/training/ProgrammaticGuidance/GenderContinuum.aspx for training materials, activities and other tools). To make a difference, we would say that anything other than gender transformative is *not* good enough. Then, calling out what is acceptable and what is not is necessary. Clarifying where interventions are at on a 'continuum' related to 'gender' *and* related to 'equity' will help to identify strategies for what needs to be done. Likewise, M&E questions (descriptive, causal, evaluative and action) need to help explore and challenge the status quo. In *Addressing gender in impact evaluation*, Fletcher (2015) provides good examples of evaluation questions that address gender in the design, delivery and assessment of interventions. For example:

- What major gender stereotypes, norms and judgements existed among intervention staff and community members at the start and at the end of the intervention in relation to the intervention topic, and, particularly, in relation to power and decision making?

- Did these stereotypes and norms shift, or become less clearly attached to judgements of worth? Did the norms become less restrictive? (Fletcher 2015:13)

Supporting meaningful engagement of the less powerful

While we still need to convince and work with donors and governments, equity-focused M&E cannot be done only with a top-down approach. We need to understand and draw on the lived experiences of those who suffer forms of inequality, inequity and injustice. But, are we paying enough attention to creating both opportunity and capacity for potential intervention beneficiaries to be meaningfully and actively engaged in intervention design, delivery and M&E processes?

Any evaluation intending to use participatory approaches needs to clarify what value this will add to the evaluation itself as well as to the people who would be involved, but also considering potential risks of their participation. (Guijt 2015) suggests that: three questions need to be answered in each situation: (1) What purpose will stakeholder participation serve? (2) Whose participation matters, when and why? and (3) When is participation feasible? The three questions are interlinked—answering them is iterative and depends on what is feasible. For example, a participatory process might seek to transform power relations, which often requires investing in building the capacity of less experienced stakeholders in evaluation and allowing more time than would be necessary if using a team of external evaluators. If these conditions cannot be met, however, then all three questions must be revisited" (Guijt 2015: 3–4).

Strengthening national M&E capacity for the long term

Evaluating the new SDGs provides a major impetus for addressing equity in M&E. In terms of strengthening national M&E capacity for data collection, analysis and reporting on the SDGs, lessons can be heeded from experience with evaluating the MDGs. In 2009, a special issue of the Journal of Acquired Immune Deficiency Syndromes (JAIDS) was devoted to progress and challenges in global data collection and reporting on MDG 6 ('halting and reversing the HIV epidemic by 2015'). These experiences are relevant to M&E of other areas of development and global health as the HIV response has been broad-based and multi-sectoral in nature (Warner-Smith et al. 2009).

Strengths and weaknesses noted in relation to investments in national M&E systems:

- global reporting catalysed the investments in basic national M&E systems but were inadequate to make these systems fully functional and sustainable;

- data availability and quality improved but major gaps remained, especially on data related to the most vulnerable or marginalised;

- insufficient attention was paid to evaluation.

The global reporting process catalysed the development of national monitoring systems and concerted M&E

strengthening efforts with substantial financial investments resulted in a marked increase in the number of countries with *basic* M&E systems and human resources devoted to M&E. However, M&E expenditures were below the minimum needed to make M&E systems fully functional (including effective data use) and the lack of long-term funding commitment by governments threatened the sustainability of the system in the long run (Peersman et al. 2009a).

Overall, both data availability and quality improved over time, but major gaps remained in data on those most vulnerable or marginalised (such as men who have sex with men, sex workers, injecting drug users, migrants). This was due to a lack of 'recognition' of certain populations by governments or, worse, criminalisation of particular behaviours (Gruskin et al. 2009) and a lack of coordination between the data collection efforts of civil society organisations working with these populations (Rugg et al. 2009).

Several critical questions about the progress and effectiveness of national responses remained unanswered as insufficient attention was paid to funding and implementing evaluation studies (Rugg et al. 2009).

Strengths and weaknesses noted in relation to increasing participation from civil society (including from 'most-at-risk populations'):

- opportunities were created for civil society to participate but there was still dissatisfaction with the 'official' reports submitted by governments;
- an important barrier for civil society participation was lack of experience working with governments and inadequate technical support.

Opportunities were created for civil society to monitor and challenge government commitments and action (Rugg et al. 2009). Civil society review of country progress reports indicated their dissatisfaction with the comprehensiveness of several reports, especially in relation to data on most-at-risk population—some governments considered such data 'irrelevant' to the epidemic. Dissatisfaction with the official reporting process and/or report was such that civil society groups in 33 countries submitted shadow reports in the 2006 reporting round and 15 countries did so in the 2008 reporting round—some of which focused on additional data considered necessary to understand the national response (Peersman et al. 2009b)

Civil society participation was strongest when civil society groups took the initiative and organised themselves but an important barrier was their lack of experience with national level processes and low access to timely technical support (Peersman et al. 2009b).

Implications for national M&E strengthening to evaluate the SDGs

Firstly, we should develop and consistently apply 'requirements' for what counts as equity-focused and gender-responsive M&E. This should include:

- explicit attention and documented plans on how gender and equity issues will be addressed across all M&E tasks;

- a comprehensive assessment of and explicit rationale for who should be involved in what M&E tasks;
- adequate resources and time and innovative ways to support meaningful and active engagement of intended beneficiaries, especially those most disadvantaged, and appropriate M&E of these;
- decision making processes that do not favour the most powerful, and appropriate M&E of these;
- going beyond assessing average effects to collect and compare results for different sub-groups or sub-areas, explicitly addressing those most disadvantaged;
- explicit definitions of 'hurdle' requirements for what is considered success in terms of gender and equity 'transformative' processes and outcomes;
- making the results of the above processes and outcomes public in a timely manner so they can contribute to cumulative learning across countries and interventions.

Secondly, we need to:

- support *long term* planning for national systems including integration of M and E, clarifying roles and responsibilities of different actors at national, sub-national and local levels of the M&E system;
- help secure long term funding;
- support for regular system assessments and prioritisation of system-strengthening activities using an asset-based approach.

References

Bamberger M, M Segone (2011). *How to Design and Manage Equity-focused Evaluations*. Evaluation Working Papers. New York: UNICEF Evaluation Office.

Bustelo M, J Espinosa, A Faúndez, M Weinstein (2015). *Guide to Including a Gender Perspective in VOPEs: Innovating to improve institutional capacities*. RELAC, EES, IOCE, EvalPartners.

Fletcher G (2015). *Addressing Gender in Impact Evaluation*. A Methods Lab publication. London: Overseas Development Institute & Melbourne: BetterEvaluation.

Gibney L, R DiClemente, S Vermund (eds) (1999). *Preventing HIV in Developing Countries: Biomedical and behavioural approaches*. New York: Kluwer/Plenum.

Gough D, S Oliver, J Thomas (2012). *An Introduction to Systematic Reviews*. London: Sage Publications Ltd.

Gruskin S, L Ferguson, G Peersman, D Rugg (2009). 'Human rights in the global response to HIV: Findings from the 2008 United Nations General Assembly Special Session Reports'. *JAIDS* 52(2):S104–S110.

Guijt I (2015). *Participatory Approaches*. Methodological Brief on Impact Evaluation, No 5. Florence: UNICEF Office of Research.

Interagency Gender Working Group (2009). *Gender Integration Continuum*. Available from: http://www.igwg.org/igwg_media/Training/FG_GendrIntegrContinuum.pdf (accessed 11 April 2016).

Mathematica Policy Research Inc (2002). *Making a Difference in the Lives of Infants and Toddlers and Their Families: The impacts of Early Head Start*. Volumes I–III: Final Technical Report [and] Appendixes [and] Local contributions to understanding the programs and their Impacts. Retrieved from http://files.eric.ed.gov/fulltext/ED472186.pdf.

Oliver S, G Peersman (eds) (2001). *Using Research for Effective Health Promotion.* Buckingham: Open University Press.

Peersman G, A Harden, S Oliver, A Oakley (1999). 'Discrepancies in findings from effectiveness reviews: The case of health promotion interventions to change cholesterol levels'. *Health Education Journal* 58:192–202.

Peersman G, D Rugg (2004). 'Intervention research and program evaluation: The need to move beyond monitoring'. In: Rugg D, G Peersman, M Carael (eds). *Global Advances in HIV/AIDS Monitoring and Evaluation. New Directions for Evaluation* 103:141–58.

Peersman G, D Rugg, T Erkkola, E Kiwango, J Yang (2009a). 'Are the investments in national HIV monitoring and evaluation systems paying off?' *JAIDS* 52(2):S87–S96.

Peersman G, L Ferguson, MA Torres, S Smith, S Gruskin (2009b). 'Increasing civil society participation in the national HIV response: The role of UNGASS reporting'. *JAIDS* 52(2):S97–S103.

Quinn Patton M (2008). *Utilization-focused Evaluation* (4th edition). Thousand Oaks: Sage Publications Inc.

Rogers P (2016). 'Understanding and supporting equity. Implications of methodological and procedural choices in equity-focused evaluations'. In Donaldson S, R Picciotto (eds). *Evaluation for an Equitable Society*. Charlotte: Information Age Publishing. Chpt 11: p. 199–213.

Rugg D, H Marais, M Carael, P De Lay, M Warner-Smith (2009). 'Are we on course for reporting on the Millenium Development Goals in 2015?' *JAIDS* 52(2):S69–S76.

Stern E, N Stame, J Mayne, K Forss, R Davies, B Befani (2012). *Broadening the Range of Designs and Methods for Impact Evaluations*. Report of a study commissioned by the Department for International Development (DFID) London: DFID, Working Paper 38.

UNICEF (2011). *Sharpening the Equity Focus. Selected innovations and lessons learned from UNICEF-assisted programmes 2009–2010. Policy and Practice*. New York: UNICEF Information and Knowledge Management Unit.

United States Institute of Peace Evaluation/USIP (2013). *Request for Proposals of USIP Projects and Grants in Iraq June 21, 2013*. Available on: http://www.usip.org/sites/default/files/IraqEvaluationsRfP.pdf.

Warner-Smith M, D Rugg, L Frescura, S Moussavi (2009). 'Monitoring the 2001 Declaration of Commitment on HIV/AIDS'. *JAIDS* 52(2):S77–S86.

Welch V, M Petticrew, J Petkovic, D Moher, E Waters, H White, P Tugwell, the PRISMA-Equity Bellagio group (2016). 'Extending the PRISMA statement to equity-focused systematic reviews (PRISMA-E 2012): Explanation and elaboration'. *Journal of Clinical Epidemiology* 70:68–89.

Westhorp G (2008). 'Development of realist evaluation models and methods for use in small-scale community based settings'. Doctoral dissertation. Nottingham: Trent University.