Systematic review

Effectiveness of interventions to strengthen national health service delivery on coverage, access, quality and equity in the use of health services in low and lower middle income countries

by Barbara Willey
Lucy Smith Paintain
Lindsay Mangham-Jefferies
Josip Car
Joanna Armstrong Schellenberg

June 2013

UKaid from the British people
London School of Hygiene & Tropical Medicine
Imperial College London
EPPI CENTRE
IOE LONDON Leading education and social research
Institute of Education University of London

This material has been funded by the Department for International Development. The views expressed do not necessarily reflect the views of the Department for International Development. The authors are part of the London School of Hygiene and Tropical Medicine and Imperial College London and were supported by the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre).

The EPPI-Centre reference number for this report is 2116.

This report should be cited as:

Willey B, Smith Paintain L, Mangham-Jefferies L, Car J, Armstrong Schellenberg J (2013) Effectiveness of interventions to strengthen national health service delivery on coverage, access, quality and equity in the use of health services in low and lower middle income countries. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

ISBN: 978-1-907345-55-5

Contents

Abbreviations

AL	Arthemeter-lumefantrine
ANC	Antenatal care
ARI	Acute respiratory infection
CBA	Controlled Before-and-after study
CI	Confidence interval
CQ	Chloroquine
CRT	Cluster randomised trial
CT	Cluster trial
DfID	UK Department for International Development
DPT3	Diphtheria, pertussis and tetanus vaccination dose 3
EPOC	Effective practice and organisation of care
ETAT+	Emergency triage, assessment, treatment and admission care
GAVI	Global Alliance for Vaccines and Immunization Alliance
HAZ	Height for age z score
HIC	High-income country
HIV	Human immunodeficiency virus
HW	Health worker
IMCI	Integrated management of childhood illness
ITS	Interrupted time series
LMIC	Low- or middle-income country
MESH	MEdical Subject Heading
PP	Per protocol
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
RCT	Randomised controlled trial
RH	Reproductive health
SMS	Short message service
SQUIRE	Standards for quality improvement reporting excellence guidelines
TEHIP	Tanzania Essential Health Interventions Project
WAZ	Weight for age z score
WHO	World Health Organization
WHZ	Weight for height z score

Executive summary

Low coverage of effective and cost-effective interventions that could save lives has been partially attributed to weak and inefficient health systems, leading to the identification and promotion of health system strengthening as a global health priority. Front-line health workers are key to delivering health services, so we have assessed the effectiveness of supply-side interventions to improve their ability to deliver health services.

Key messages

Moving beyond technical guidance alone:

- Studies which strengthened other elements of the health service delivery in addition to technical guidance, as well as community mobilisation and interventions at the health sector policy and strategic management level showed more consistent improvement on quality of care and counselling than those using technical guidance alone.

Supply-side interventions that appeared to have a positive effect on quality of care included:

- text message reminders (with motivational quotes) for malaria case management;
- training for malaria case management when combined with community awareness, supervision and referral mechanisms;
- job aids for antenatal counselling when combined with supervision and a focus on institutional adaptations required to incorporate the use of these job aids;
- IMCI (integrated management of childhood illnesses) training, when implemented in combination with enhanced supervision that incorporated training of supervisors, job aids, use of data and face-to-face supportive supervision, in repeated cycles of assessment, examination/feedback and planning;
- implementation of guidelines, when delivered using training, enhanced supportive supervision, a focal person to troubleshoot problems on site, and repeated progress surveys with face-to-face feedback and planning sessions;
- quality improvement, when combined with training, supervision, repeated progress surveys with time frames and named individuals identified against decisions/plans made during face-to-face meetings with all health facility staff, and district level representation;
- implementation of full IMCI guidelines, incorporating training, supervision and discussion of how to overcome barriers to implementation, wider health system strengthening at the health sector policy and strategic management level, and community mobilisation.

Background

About the systematic review underlying this summary

Review objectives:

1. to identify and describe characteristics of supply-side interventions that are intended to improve health services provided by front-line workers;

2. to assess and report the effectiveness of these interventions on:

    a. coverage of health services;

    b. access to health services;

    c. quality of health services;

    d. equity in the use of health services; and

    e. morbidity and mortality.

	What the review authors searched for	What the review authors found
Interventions	Supply-side interventions to improve the health services provided by front-line workers during their interaction with users. We included interventions to improve the delivery of existing services; we did not include interventions to test the delivery of services through new cadres or health workers, or the delivery of new services.	A wide variety of interventions were documented. Three categories of intervention were identified and described, lying along a spectrum of increasing incorporation of other elements of the health service delivery level and/or other levels at which constraints to improving access to health care act.
Participants	Front-line health workers employed by national governments to delivery services at primary, secondary or tertiary facilities, including community health workers.	The majority of studies included health workers managing children at first level or referral level facilities, such as doctors, clinical officers, nurses and in some cases nursing aides.
Settings	Low and lower middle income countries. Delivery in at least one district.	12 studies were conducted in 9 countries - Bangladesh (1), Benin (2), Guinea (1), Kenya (3), Mali (2), Mexico (1), Pakistan (1), Thailand (1), Tanzania (2). 7/12 were in rural settings. 8/12 covered more than one district.

Outcomes	Coverage Access (physical access to healthcare) Quality (process of care) Equity of coverage, access or quality Survival impact - under-five mortality	Most studies reported multiple effect measures and many did not specify a primary outcome.
Date of most recent search: January, 2012		
Limitations: The search strategy devised did not perform well in identifying grey literature, which could introduce bias. This may have an impact on the completeness of the findings of this review in relation to the first objective of describing and characterising interventions to improve health service delivery. Due to the broad range of included interventions and outcomes, meta-analysis was not appropriate. Data were aggregated instead based on narrative synthesis using a thematic summary approach (Thomas et al., 2012). The broad nature of the systematic review question and the focus on effectiveness of interventions did not enable detailed logic frameworks for each intervention that identified and illustrated interim processes and outputs between intervention inputs and assessed outcomes.		

We conducted a systematic review to synthesise the evidence for the effectiveness of interventions to strengthen national health service delivery on coverage, access, quality and equity in the use of health services in low and lower middle income countries. Health system strengthening includes many components and can occur at many levels. Two frameworks are used to position and structure this review: the WHO's six building blocks of health system components (World Health Organization, 2007) and a framework of five levels at which constraints to improved access to health care exist: I. community/household; II. health services delivery; III. health sector policy and strategic management; IV. public policies cutting across sectors; and V. environmental and contextual (Hanson et al., 2003).

Summary of findings

1. Characterisation of interventions

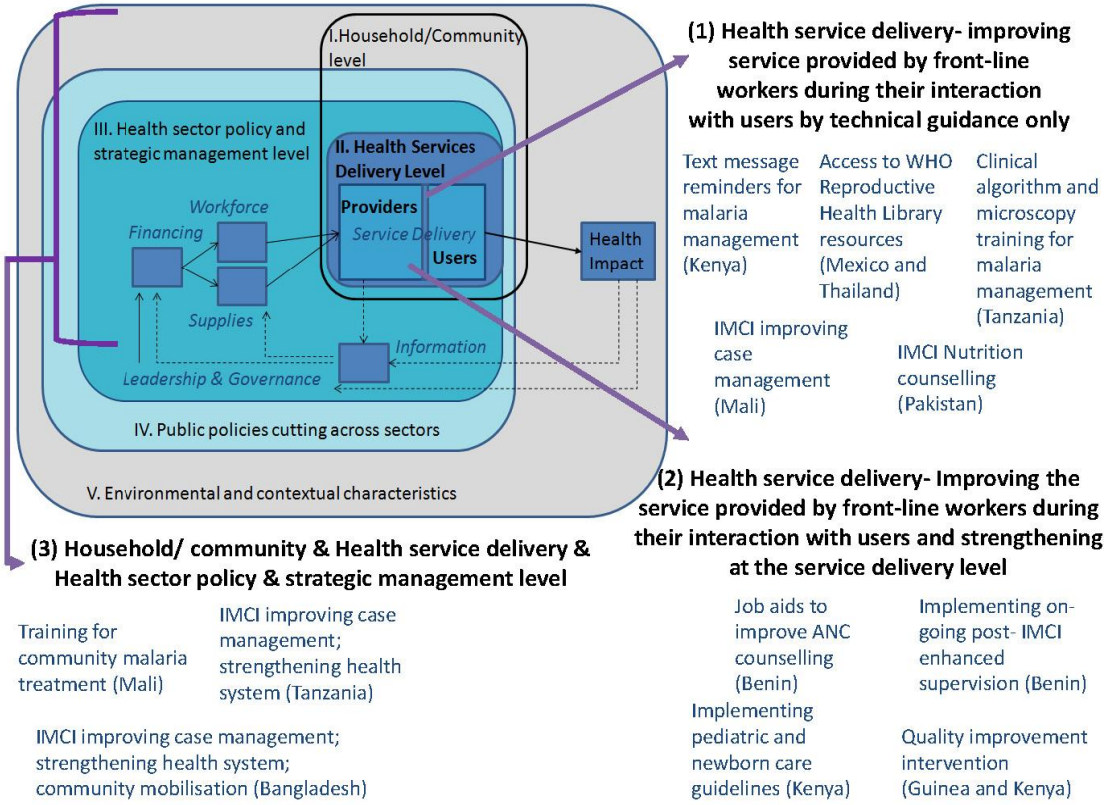
Findings from this review suggest that studies which strengthened other elements of the health service delivery in addition to technical guidance, as well as community mobilisation and interventions at the health sector policy and strategic management level, showed more consistent improvement on quality of care and counselling, than those using technical guidance alone.

- Five studies were classified in the first category (A), which principally included training, although one study also included text message reminders and one included access to electronic resources summarising best practice.

Effectiveness of interventions to strengthen national health service delivery on coverage, access, quality and equity in the use of health services in low and lower middle income countries

- Four studies were classified in the second category (B), and included: training; frequent supervision with face-to-face feedback and supervision of supervisors; cycles of quality improvement and assessment with feedback on progress; the creation of on-site teams to identify changes required to implement the intervention, and nomination of a full time on-site facilitator to encourage implementation of the intervention.

- Three studies were classified in the third category (C), and included: training; introduction of job aids and sensitisation of staff at referral centres to encourage referral; introduction of an improved supply tracking and reporting system at facility level; introduction of improved management, budget and planning systems at district level.



2. Quality of studies and effectiveness of interventions

Risk of bias was assessed using the EPOC checklist (see Appendix 7). Ten studies used a cluster randomised trial (CRT) design, and two a non-randomised quasi-experimental cluster design; both, as expected, were categorised as being of high risk of bias. Of the ten CRTs, only one was at low risk of bias across all categories, with an additional four at low risk of bias across the majority of categories.

Most studies reported follow-up or presented trial profiles to allow determination of whether there was incomplete outcome assessment.

Where this was available, follow-up was generally above 80 percent. However, some studies did not present this information and are therefore graded 'not clear'.

On the whole, blinded assessment of the main study outcome was not carried out. However, as the outcome was frequently quality of care and assessed using

structured survey tools or checklists, outcome assessment is unlikely to have been subjective, reducing potential measurement bias.

Similarity between groups at baseline was reported with disappointingly low frequency, making it hard to assess success of randomisation, which may be a particular issue in cluster designs.

Quality of studies included in the review

Reference	Outcome				Contamination	Randomisation	Incomplete outcome data	Blinding of outcome assessment	Similarity at baseline	Power
	Coverage	Morbidity Mortality	Quality	Equity						
Arifeen et al. (2004a, b)	X	X	X		✓	✓	✓	✗	✓	✓
Armstrong Schellenberg et al. (2004a, b)	X		X	X	✓	✗	✓	✗	✗	✗
Ayieko et al. (2011)	X		X		✓	✓	✓	✗	✓	✓
Bradley and Igras (2005)			X		✓	✗	✓	✗	?	✗
Gilroy et al. (2004)			X		✓	✓	?	?	✗	✓
Gulmezoglu et al. (2007)			X		✓	✓	✓	✗	✓	✓
Jennings et al. (2010)			X		✓	✓	✓	✗	✗	✓
Ngasala et al. (2008)	X				✓	✓	✓	✗	✗	✓
Rowe et al. (2009)			X		✓	✓	✗	✗	?	?
Winch et al. (2003)	X		X		✓	✓	?	✗	✗	✓
Zaman et al. (2008)		X	X		✓	✓	✓	✓	✓	✓
Zurovac et al. (2011)			X		✓	✓	✓	✗	✓	✓

It was not appropriate to pool effect estimates in a meta-analysis. The effectiveness of interventions on coverage, quality, equity and impact are discussed in each of the three classifications of intervention.

Coverage

Five studies reported on nine coverage outcomes, including health seeking, breastfeeding, and management of childhood diseases.

Three studies reported on two outcomes of clinical impact: anthropometric outcomes (three studies) and under-five mortality (two studies).

Of five studies reporting coverage, two (one each from categories A and B) reported significant improvements of 19-35 percentage points in outcomes of malaria treatment prescriptions, uptake of HIV testing and recording of child's immunisation status. Three studies (classified in category C) reported data on outcomes that were closely linked to morbidity and mortality. In these, the authors recorded significant improvements: of 10 percentage points in exclusive breastfeeding; 6-10 percentage points in appropriate health-seeking behaviour; 55-63 percentage points in the appropriate management of malaria; 35 percentage points in the appropriate management of pneumonia; and 58 percentage points in the management of priority illnesses in children.

In terms of under-five mortality, two studies both of integrated management of childhood illness (IMCI), reported a non-significant improvement in under-five mortality of 27 per 1,000 live births, in comparison to a rate of 28.2 per 1,000 and 31.2 per 1,000 in comparison areas, in Tanzania and Bangladesh respectively.

Access

None of the 12 studies reported outcomes of access.

Quality of care

The majority (11/12) reported quality outcomes. In terms of studies from category A, one showed mixed improvement in some aspects of nutritional counselling for children, while a second reported significant improvement of about 24 percentage points in the quality of malaria treatment in children. Findings from all studies classified within categories B and C showed consistently positive significant improvements. Although variation in methods of measuring quality don't allow simple numerical summaries, those in category B showed significant improvements of 7-19 percentage points in mean scores, while in category C, the IMCI evaluation from Bangladesh reported an improvement of 56 percentage points in mean scores of quality of case management of childhood infections.

Equity

Only one study, the Tanzanian evaluation of IMCI, reported outcomes of equity. The results showed a positive association between strengthening health service delivery and equity of treatment of underweight, stunting and fever, as well as measles vaccination and bednet use, using a concentration index of household-level asset ownership, household characteristics, education and income. Although improvements in equity were seen, malnutrition outcomes remained pro-rich, with more children in the poorest quintile compared to the least poor quintile classified as malnourished.

3. Potential pathways of effect and mapping the evidence base

The evidence base for quality outcomes is the largest, with the majority of studies reporting this outcome (green line).

This review found no or sparse evidence of interventions delivered at scale and evaluated robustly for outcomes of access or equity (red lines).

The findings from the review show some mixed evidence on a wide range of coverage indicators, which vary in how closely they relate to measures of disease prevalence or incidence (amber lines).

4. Shared characteristics of success

Findings from synthesising the effectiveness evidence show that of the eleven studies that reported quality outcomes, eight showed substantial and consistent improvement following the intervention. Within these eight studies, we identified some shared characteristics relating to the interventions, based on the published descriptions of the interventions. These have been broadly summarised into four characteristics:

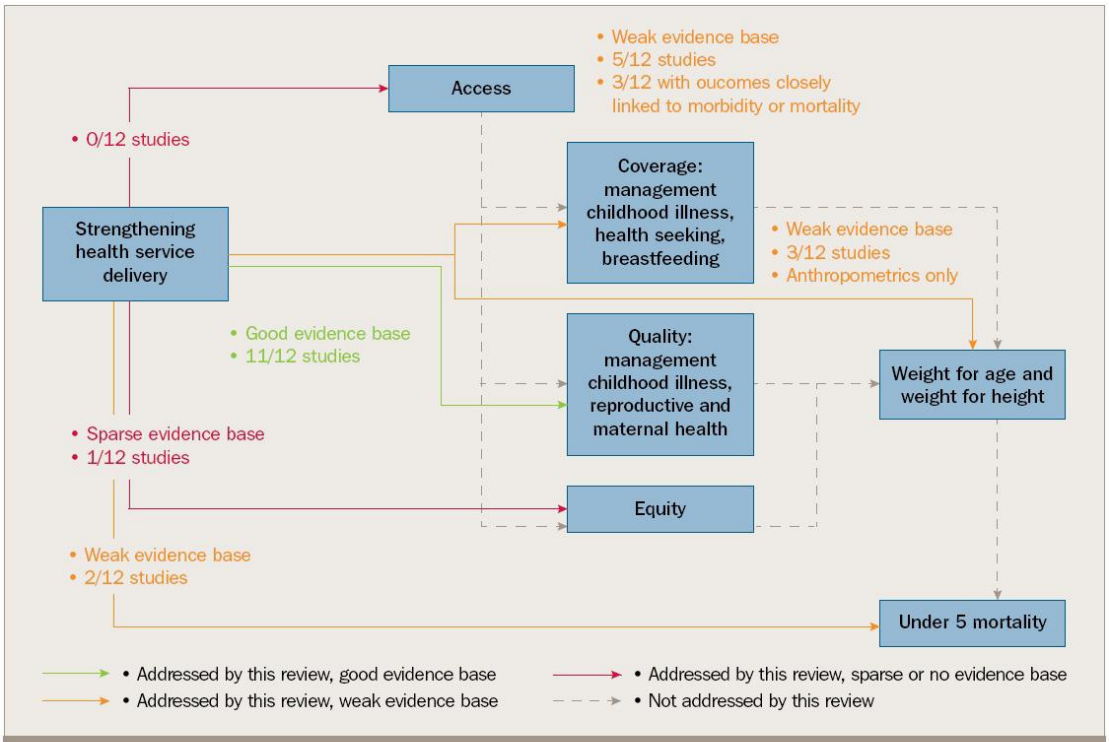
Sustained interaction: studies with repeated interactions between the intervention and the health worker, and those that had continued interaction over a substantial period tended to show more consistently positive findings.

Local ownership through a focal point person: many of the studies reporting successful outcomes incorporated an individual, often an existing health facility staff member, to act as the on-site point person for the intervention, which seemed to be important in fostering ownership of the intervention by participating facilities.

Feedback cycles: studies that incorporated feedback to health staff, often through continued supportive supervision or problem solving, although sometimes through interim monitoring and assessments, showed more consistently positive findings.

Beyond the health facility: many of the studies reporting successful quality outcomes included aspects of the intervention that went beyond clinical management and targeted management, supervision and wider systems management (e.g. record keeping, supplies monitoring) at the health facility level, but also linked these processes and their findings to management and planning systems beyond the health facility, for example to the district.

Effectiveness of interventions to strengthen national health service delivery on coverage, access, quality and equity in the use of health services in low and lower middle income countries



Mapping the availability of evidence to the logic framework: linking strengthening health service delivery and mortality

Additional information

Related literature

We identified 12 systematic reviews (one of which is still in process) related to interventions implemented to target constraints at the health service delivery level. Eight relate directly to interventions to address weak technical guidance, programme management and supervision of health workers in LMICs: Bosch-Capblanch et al. (2011), Forsetlund et al. (2009), Jamtvedt et al. (2003), Oliveira-Cruz et al. (2003), Opiyo and English (2010), Rowe et al. (2012), Smith et al. (2009) and van Lonkhuijzen et al. (2010).

We can make a number of observations based on the evidence synthesised in this review, and on that available from other systematic reviews on linked topics:

Evidence base for efficacy and effectiveness: This review, and others, have emphasised the relative scarcity of studies carried out using robust study designs to generate evidence of efficacy or effectiveness in LMICs. Furthermore, other reviews which undertook assessment of the quality of the evidence for impact (e.g. lay health workers) concluded that the quality of the evidence for impact is generally low to very low, as defined by the GRADE criteria (Guyatt GH, Oxman AD, Vist G, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, for the GRADE Working Group). Other reviews which had more inclusive inclusion criteria for study design, commented on the difficult of comparing studies due to poor study quality and variability in methods of measuring outcomes.

Outcomes: There is a focus on outcomes of health worker knowledge, with fewer assessing health worker behaviour and very few assessing outcomes of survival impact.

Technical guidance alone is insufficient: Although the authors of most of the other reviews identified did not explicitly classify interventions by their

components, other reviews did emphasise that simple training interventions often did not have positive or large effects on behaviour, and that interventions such as educational meetings alone were insufficient to change complex behaviour. The authors of other reviews highlighted that training was more effective when included as part of a package of interventions, combined with, for example, enhanced supportive supervision.

Shared characteristics of success: Our review also highlighted three other characteristics in addition to moving beyond technical guidance alone. These included: 1) sustained interaction, 2) local ownership through a focal point person and 3) feedback cycles. These findings were not specifically drawn out in the results or conclusions of other systematic reviews.

1. Background

1.1 Missed opportunities

- 6.9 million children under five years of age, including almost 3 million neonates, died in 2011
- Over 65 percent of these deaths in under-fives and up to 55 percent of those in neonates could be avoided if known effective interventions were delivered at scale
- The missed opportunities caused by weak health systems failing to deliver such interventions are increasingly recognised

In 2011, there were 6.9 million deaths in children under the age of five. Forty percent of these occurred during the first seven days of life (UNICEF, 2012, World Health Organization, 2009, 2010b). Low and middle income countries (LMICs) experience 99 percent of child deaths (World Health Organization, 2009). Leading causes of death in LMICs remain dominated by infectious disease. These include respiratory infections, diarrhoea, HIV, malaria, tuberculosis and neonatal-related complications, such as prematurity, intrapartum-related deaths or birth asphyxia, and neonatal infections.

Two-thirds of child deaths, and between 35 and 55 percent of neonatal deaths, could be avoided by implementing known effective and cost-effective interventions at scale (Claeson et al., 2003, Darmstadt et al., 2005, Jones et al., 2003). Failure to do this has been partially attributed to weak and inefficient health systems (Travis et al., 2004). This has led to the identification and promotion of health system strengthening as a global health priority (Bryce et al., 2003, de Savigny and Adam, 2009, Frenk, 2010, van Olmen, 2010, 2012b, World Health Organization, 2007). Renewed efforts to strengthen health systems, including health service delivery, have been seen (Bryce et al., 2003, Frenk, 2010, Fryatt et al., 2010, Madon et al., 2007, Travis et al., 2004). Increased focus from donors to explicitly encourage the inclusion of health system strengthening interventions in grant applications has been seen in recent years (e.g. GAVI and the Global Fund). Several international initiatives dedicated to strengthening health systems have also been established (e.g. the Implementation Research Platform hosted by the Alliance for Health Policy and Systems Research, the International Health Partnership (IHP+), and the High-level Taskforce on Innovative Financing for Health Systems (Bennett et al., 2008).

The challenge for global health is to translate these efficacious interventions into effective public health policies that are successfully implemented at scale. The emphasis of evaluating interventions to improve health delivered at scale has been highlighted by *The Lancet*, which emphasized that effectiveness evaluations of large-scale global health programmes 'must now become the top priority in global health' (Lancet, 2010). Evaluations of interventions implemented under near-programmatic conditions, with reported detail on context, are necessary to aid understanding of why and how interventions are effective, provide evidence on implementation, and inform policy makers in other settings to enable them to establish whether the intervention and its outcomes are reproducible in their setting.

1.2 Health systems and constraints to the delivery of health services

- Health systems are complex, context-specific and adaptive systems
- We have used the WHO building blocks framework to define inputs to the health system
- We have used a framework by Hanson and colleagues to define and understand constraints to delivering health services

The mechanisms through which health system strengthening interventions are anticipated to result in improved health and reduced mortality are complex, and constraints to improved access to health care exist at different levels (de Savigny and Adam, 2009, Hanson et al., 2003).

In this review, we have drawn on two conceptual frameworks (World Health Organization and Hanson et al.) that are used:

- to aid definition and understanding of health systems;
- to identify constraints to improved delivery of health services; and
- to identify interventions to overcome constraints to improved delivery of health services.

Frameworks and definitions of health system strengthening (Remme et al., 2010) and health system performance (Murray and Frenk, 2000) are not wholly set. However, the WHOs six building blocks are a widely used starting point to define health system inputs (World Health Organization, 2007) (Figure 1.1).

The suggested building blocks are:

1. service delivery
2. workforce (human resources)
3. information
4. medicines and technologies
5. financing
6. leadership and governance

Recently, the building blocks have been presented as six overlapping circles with people at the centre.

Figure 1.1: WHO Health System Framework (de Savigny and Adam, 2009)

This presentation of the building blocks also highlights that people, and in the case of service delivery, front-line workers, are central.

This helps to emphasise that the building blocks represent inputs and that it is important to additionally consider the interrelations between blocks and apply a systems perspective (de Savigny and Adam, 2009).

A framework by Hanson and colleagues (Hanson et al., 2003) highlights five levels at which constraints may act (Table 1.1). These levels are:

I. community and household;

II. health services delivery;

III. health sector policy and strategic management;

IV. public policies cutting across sectors; and

V. environmental and contextual characteristics.

Table 1.1: Levels of constraints to improving access to priority health care, reproduced from Hanson et al. (2003)

Level of constraint	Types of constraint
I. Community and household level	Lack of demand for effective interventions Barriers to use of effective interventions (physical, financial, social)
II. Health services delivery level	Shortage and distribution of appropriately qualified staff Weak technical guidance, programme management and supervision Inadequate drugs and medical supplies Lack of equipment and infrastructure, including poor accessibility of health services
III. Health sector policy and strategic management level	Weak and overly centralised systems for planning and management Weak drug policies and supply system Inadequate regulation of pharmaceutical and private sectors and improper industry practices Lack of intersectoral action and partnership for health between government and civil society Weak incentives to use inputs efficiently and respond to user needs and preferences Reliance on donor funding that reduces flexibility and ownership Donor practices that damage country policies
IV. Public policies cutting across sectors	Government bureaucracy (civil service rules and remuneration; centralised management system; civil service reforms) Poor availability of communication and transport infrastructure
V. Environmental and contextual characteristics	Governance and overall policy framework: • Corruption, weak government, weak rule of law and enforceability of contracts • Political instability and insecurity • Low priority attached to social sectors • Weak structures for public accountability • Lack of free press Physical environment: • Climatic and geographic predisposition to disease • Physical environment unfavourable to service delivery

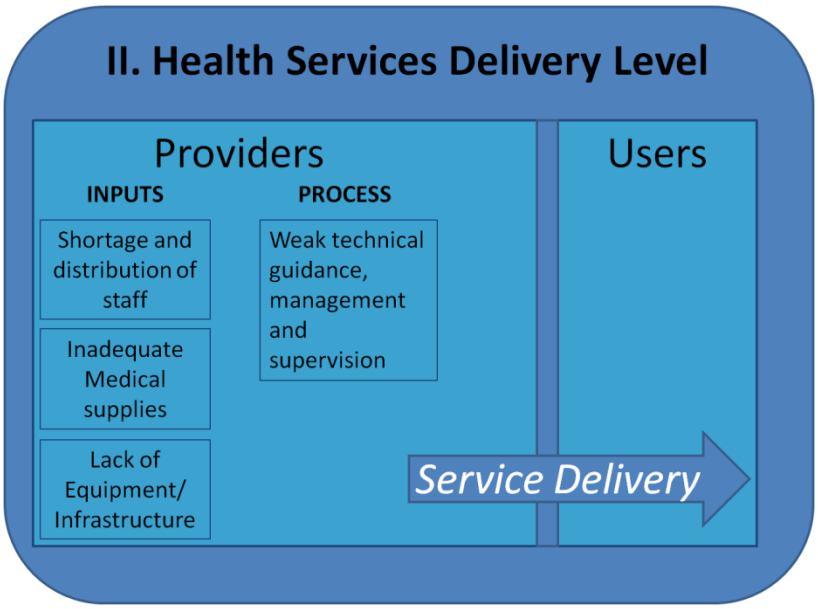
Hanson et al. make a useful distinction between factors within a community or household that may affect the demand for health care, and factors that exist at the service delivery level that may affect its supply. They specifically identify four constraints acting at the health service delivery level.

- shortage and distribution of appropriately qualified staff;
- weak technical guidance, programme management and supervision;
- inadequate drugs and medical supplies;
- lack of equipment and infrastructure, including poor accessibility of health services.

The shortage and distribution of staff, inadequate supplies and lack of equipment and infrastructure can be thought of as inputs. The constraints of weak technical guidance, programme management and supervision can be thought of as processes linking the inputs of staff, supplies and infrastructure to service delivery (Figure 1.2).

Addressing constraints relating to inputs are unlikely to improve service delivery in the absence of either the capacity by front-line workers to deliver services to a defined standard or quality or appropriate supervision and management of front-line workers.

Figure 1.2: Constraints acting at the health service delivery level (Hanson et al., 2003)



1.3 Reviewing the evidence

- This systematic review is located within a wider body of evidence synthesis for health system strengthening
- Attention is drawn to existing reviews addressing constraints at the community/ household level and at the health sector policy and strategic management level
- Attention is drawn to existing reviews that address constraints at the health service delivery level
- This review aims to fill a gap in the evidence synthesis of approaches that address constraints at the health service delivery level by focusing on interventions delivered at scale in low and lower middle income countries

A considerable body of evidence synthesis for health system strengthening exists, particularly at the community and household level and the health sector policy and strategic management levels of Hanson and colleagues' framework (levels I and III shown in Table 1.1).

At the community and household level, a number of reviews have been completed, including:

- user fees (Lagarde and Palmer, 2011);
- demand-side financing (Lagarde et al., 2007, Lagarde et al., 2009); and
- user-side interventions to improve malaria treatment (Smith et al., 2009).

At the health sector policy and strategic management level, reviews exist on:

- integration of health services (Briggs and Garner, 2006, Briggs et al., 2001, Dudley and Garner, 2011);
- incentives (Eldridge and Palmer, 2009) and pay-for-performance interventions for health workers in LMICs (Witter et al., 2012);
- contracting out health services (Lagarde and Palmer, 2009); and
- health insurance (Spaan et al., 2012).

The evidence for improving health service delivery by addressing the four constraints identified by Hanson and colleagues (shortage and distribution of staff; inadequate medical supplies; lack of equipment and infrastructure; and weak technical guidance, programme management and supervision) has to some extent been previously synthesised.

Reviews have focused for the most part on the shortage and distribution of staff, including strategies for task-shifting to lay or community health workers, and interventions targeting weak technical guidance, weak programme management and weak supervision. Constraints relating to medical supplies and infrastructure have often been addressed as components of interventions to address these two sets of constraints. Some evidence synthesis has focused on delivery of health services in general, while many have taken a disease-specific focus, particularly for malaria, and newborn and maternal health (Table 1.2).

Particularly relevant to this review, Alexander Rowe and colleagues have been working for the past six years to comprehensively synthesise the evidence on interventions to improve health worker performance in LMICs (Rowe et al., 2012). Preliminary results suggest a substantial evidence base (~500 studies included). However a focus on interventions delivered at scale has not been applied.

Table 1.2: Existing reviews synthesising the evidence of interventions addressing constraints to health service delivery

Review topic	Setting	Constraint addressed	Reference
General			
Interventions to reduce emigration of health care professionals from low- and middle-income countries	LMIC	Shortage and distribution of appropriately qualified staff	Penaloza et al. (2011)
Interventions for increasing the proportion of health professionals practising in underserved communities	LMIC	Shortage and distribution of appropriately qualified staff	Grobler et al. (2009)
Continuing education meetings and workshops: effects on professional practice and health care outcomes	HICs and LMICs	Weak technical guidance	Forsetlund et al. (2009)
Audit and feedback: effects on professional practice and health care outcomes	High and LMICs	Weak programme management	Jamvedt et al. (2003)
Managerial supervision to improve primary health care in low- and middle-income countries	LMIC	Weak supervision	Bosch-Capblanch et al. (2011)
Approaches to overcoming constraints to effective health service delivery	LMIC	Weak technical guidance, programme management and supervision	Olivera-Cruz et al. (2003)
How can we achieve and maintain high-quality performance of health workers in low-resource settings?	LMIC	Weak technical guidance, programme management and supervision	Rowe et al. (2012)
Disease- and condition-specific			
Traditional birth attendant training for improving health behaviours and pregnancy outcomes	LMIC	Shortage and distribution of appropriately qualified staff	Sibley et al. (2007)

Review topic	Setting	Constraint addressed	Reference
Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases	HICs and LMICs	Shortage and distribution of appropriately qualified staff	Lewin et al. (2010)
In-service training for health professionals to improve care of the seriously ill newborn or child in low and middle-income countries	LMIC	Weak technical guidance	Opiyo and English (2010)
Interventions to improve quality of emergency obstetric care	LMIC	Weak technical guidance	van Lonkhuijzen et al. (2010)
Provider-side interventions to improve malaria treatment	LMIC	Weak technical guidance, programme management and supervision	Smith et al. (2009)

1.4: Rationale for review

There has not previously been a focus on collating evidence on the effectiveness of interventions to strengthen health service delivery implemented at scale.

It is important to assess this evidence as large-scale implementation is required to achieve high coverage, without which substantial impact on mortality would not be possible.

Additionally there are implementation challenges with large-scale deployment that may not be encountered during pilot or smaller-scale studies.

There is a need to inform policy makers and the policy decision-making process by characterising the alternative approaches that could be deployed to improve the delivery of health services by front-line workers, and by synthesising the evidence on the effectiveness of these approaches.

In addressing these needs, it is important to use information from studies that have a robust study design, to minimise the role of bias and chance in the findings, and increase the internal validity of the results. However, it is also important to address questions of implementation, as this experience is likely to be of use in policy-making decisions.

1.5 Aims for review

This review aims to assess the published and grey literature evidence for the effectiveness of supply-side interventions that are intended to improve the health services provided by front-line workers during their interaction with users.

We focus on interventions in low or lower middle income countries that are implemented and evaluated at scale, defined as an intervention implemented in at least one district (lowest level of health administration).

1.6 Specific objectives

1. To identify and describe characteristics of interventions that are intended to improve health services provided by front-line workers;

2. To assess and report the effectiveness of these interventions on:

    a. coverage of health services;

    b. access to health services;

    c. quality of health services;

    d. equity in the use of health services; and

    e. morbidity and mortality.

2. Methods

- Criteria for considering studies for this review were categorised under participants, interventions, study designs and outcomes
- We searched MEDLINE, CENTRAL and Embase electronic databases using a combination of broad search terms relating to health systems AND developing countries AND study design
- Screening and data extraction were carried out by two review authors
- Quality of included studies was assessed using the EPOC Risk of Bias checklist
- Estimating a pooled effect of effectiveness through meta-analysis was not appropriate
- Narrative synthesis was instead used to classify interventions, report effectiveness and systematically compare characteristics across interventions

2.1 Criteria for considering studies for this review

2.1.1 Types of participants

- Front-line workers employed by national governments to delivery services at primary, secondary or tertiary facilities, including community health workers, were eligible for inclusion. This is because health services in many low or lower middle income settings are principally delivered at scale by state providers.
- Front-line workers were eligible both as individual health workers and as part of wider teams (e.g. all staff from a whole health facility).
- We focused on countries defined by the World Bank as low and lower middle income economies (listed in the search strategy in Appendix 2).

2.1.2 Types of interventions

We included supply-side interventions to improve the health services provided by front-line workers during their interaction with users. We included interventions to improve the delivery of existing services, but we did not include interventions to test the delivery of services through new cadres or health workers, or the delivery of new services.

Examples of eligible interventions aimed to improve the knowledge and skills of front-line health workers

- Pre-service training
- In-service training
- Supervision
- Guideline and protocol dissemination
- Reminders
- Quality improvement
- Quality assurance
- Audit and feedback
- Checklists

Packages of interventions aiming to improve service delivery were eligible, as long as a component of the package was interventions to improve service delivery through addressing weak technical guidance, programme management and supervision.

Packages of interventions to strengthen health systems that were targeted at levels other than the service delivery level (i.e. community or household level; health sector policy and strategic management level; public policies cutting across sectors; or environmental and contextual characteristics), were also included if they also targeted weak technical guidance, programme management and supervision constraints at the health services delivery level (Hanson et al., 2003).

This review was restricted to interventions that were delivered at scale. Specifically, we included interventions that were implemented in at least one district (lowest level of health administration), where the comparison group(s) were at least one other district, or where the intervention was delivered in one district only and comparison and intervention areas or groups were at the sub-district level.

We did not include the following interventions targeting constraints at the household or community level, or at the health sector policy and strategic management level:

- Contracting
- User fees
- Demand side financing
- Demand generation
- Integration of health services
- Incentives
- Pay-for-performance interventions
- Insurance

2.1.3 Types of study designs

We included the study designs shown in the box below and in Appendix 1.2. Inclusion of these study designs was guided by Cochrane Effective Practice and Organisation of Care (EPOC) guidelines (Higgins and Green, 2011).

The inclusion of non-randomised designs was important in this review because of its focus on interventions delivered at scale; randomising large units such as whole administrative areas is frequently impractical, and individual randomisation has risks of contamination. Furthermore, closely controlled designs may have low external validity, i.e. generalisability, and there is a need to include study designs other than randomised controlled trials when assessing complex public health interventions at scale (Victora et al., 2011).

We included comparison groups that included usual or standard care, or an alternative strategy to improve health service delivery; this included comparisons of multi- versus single-strategy interventions.

Included study designs

- Individually randomised controlled trials (RCT)
- Cluster randomised trials (CRT), including randomised stepped wedge designs
- Non-randomised cluster trials (CT)
- Controlled before-and-after studies (CBA)
- Interrupted time series studies (ITS)

2.1.4 Types of outcome measures

The primary outcomes of preference to assess the effectiveness of interventions to improve health service delivery were measures of survival impact, such as under-five mortality rate or infant mortality rate. We anticipated that many studies would not have measured survival impact, and therefore we also included outcomes of coverage, access and equity. These were based on the WHO's 2011 Indicator compendium,[1] and were selected depending on availability across studies. To be included, studies needed to report at least one of the outcomes listed in the box below.

Coverage

Coverage of evidence-based interventions, that where possible include evidence of biologically-plausible mechanisms of effect on survival impact, for example:

- coverage of DPT3 vaccination;
- proportion of deliveries with a skilled attendant;
- treatment of children under five with parasite-confirmed malaria with an appropriate anti-malarial within 24 hours of onset of fever.

Access

Access is a multi-dimensional concept as defined by McIntyre and colleagues (McIntyre et al., 2009).

In this review we focus on the physical access to health services, for example:

- access to a health facility within 5km;
- ratio of health professionals to population;
- availability of specific services, such as the WHO's Essential Health Service Package.

[1] http://www.who.int/gho/publications/world_health_statistics/WHS2011_IndicatorCompendium.pdf

Effectiveness of interventions to strengthen national health service delivery on coverage, access, quality and equity in the use of health services in low and lower middle income countries

Quality We focus on the process of care, as defined by Donabedian[2] (Donabedian, 2003). We define quality in the process of care as the technical or interpersonal activities provided to a defined standard. We include outcomes of quality that may be measured by an index or proportion of essential tasks completed. We have not included user satisfaction as a measure of quality within this review, since this falls within the community-level constraints of Hanson and colleagues' framework rather than the service delivery level.
Equity Equity of included outcomes of coverage, access, or quality were assessed by disaggregating coverage by any of the following: • wealth quintiles; • education; • urban/rural residence; • gender.
Survival impact Under-five mortality; Infant mortality.

2.2 Search strategy for identifying studies

2.2.1 Electronic searches

We searched MEDLINE, CENTRAL and Embase electronic databases using a combination of broad search terms relating to health systems (health service delivery, health services, health workforce, quality assurance) AND developing countries AND study design. Complete search strategies are shown in Appendices 2 and 3. The final search strategy was translated from MEDLINE to CENTRAL and Embase databases.

In order to access grey literature, we translated the MEDLINE search strategy to the Global Health database, and browsed studies listed under the health service delivery category within ELDIS.

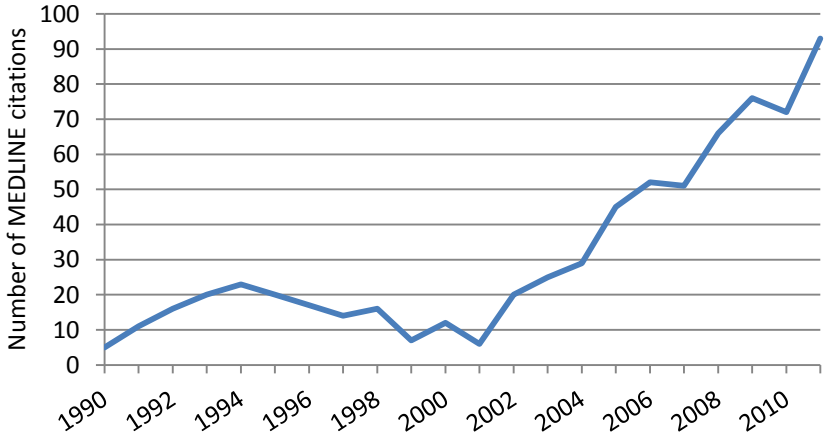
We define the papers included in this review that reported on the quantitative outcomes of intervention studies as index papers. Additionally, reference lists and citation searches on the index papers were used to identify satellite papers, and as such, it was possible to include more than one paper per study. Satellite papers were defined as publications from the same study as the index paper that reported qualitative data, e.g. on implementation, inputs, processes, outputs and outcomes that might be found along the spectrum of the hypothesised pathway of influence

[2] The Donabedian model is based on a three-component approach to assessing quality of care: structure, process and outcomes. Process of care denotes what is actually done to the patient in the giving and receiving of care.

between intervention and impact; these were included in the analysis to extract this additional information that was not found in the index paper, but which would be relevant for policy makers.

Health systems and contexts in many low and lower middle income countries have undergone substantial changes in recent times. Additionally, over the past decade or so, there has been increased interest and attention given to scaling up (Mangham and Hanson, 2010). Figure 2.1 shows a line graph of the increase in citations within MEDLINE over the period 1990-2011 for 'scaling up' terms combined with 'developing country' terms. Therefore we limited the review to literature published between 2000 and 2011 to focus on synthesising very recent experience. No limits on language of publication were applied.

Figure 2.1: Increase in MEDLINE citations for 'scaling up' AND 'developing countries' 1990-2011



The results from database and other searches were downloaded and managed within EndNote, where duplicate records were discarded.

2.3 Data collection

2.3.1 Screening

Two review authors (BW and LSP) independently assessed all the potential studies identified as a result of the search strategy against the inclusion criteria (Appendix 1.2), using an associated screening template (Appendix 4). Discrepancies in the selection of studies that could not be reconciled by discussion were resolved by referring to the full text, or as a final stage by review from a third review author (LMJ).

We described reasons for exclusion (Appendix 6), listed excluded studies, and use a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart to describe the study selection process (Higgins and Green, 2011).

2.3.2 Data extraction and management

Two review authors (BW and LSP) extracted the data into a standardised extraction form in an Access database. Categories of data extracted are listed in detail in Appendix 5, but briefly included details of the study design, intervention, results, study quality and context.

2.4 Data analysis and synthesis

2.4.1 Quality of studies: assessment of risk of bias

Two review authors (BW and LSP) independently assessed risk of bias for each included study based on the criteria outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* and the Cochrane Effective Practice and Organisation of Care (EPOC) checklist (Appendix 7).

We resolved disagreement by discussion, or by including a third review author (JS). Areas assessed are shown in the box below and each domain was assessed as 'done', 'not done' or 'not clear' suggesting low, high or unclear risk of bias respectively (Higgins and Green, 2011).

- Protection against contamination
- Random sequence generation and allocation concealment
- Blinding of main outcome assessment
- Incomplete reporting of outcome data
- Similarity of outcome and other characteristics at baseline between intervention and comparison groups
- Reporting of appropriate sample size calculation

2.4.2 Narrative synthesis

Included studies varied too greatly to make meta-analysis appropriate, particularly in terms of interventions, comparison groups and outcomes. Therefore a pooled effect using meta-analysis was not estimated (Higgins and Green, 2011). Instead, in line with recommendations for systematic reviews of complex interventions (Petticrew and Roberts, 2006), we structured the analysis and synthesis of this review by using a narrative synthesis approach of thematic summaries (Popay et al., 2006, Thomas et al., 2012).

Narrative synthesis was used to synthesise included studies to meet the objective of characterising interventions according to the Hanson et al. (2003) framework. As outlined fully in the results, interventions were put into three categories, shown in the box below.

Supply-side interventions aimed to address the constraint of:

A. weak technical guidance without wider health system involvement;
B. weak technical guidance, along with strengthening of other elements at the service delivery level:
    - programme management or supervision;
    - shortage and distribution of appropriately qualified staff;
    - inadequate drugs and medical supplies; and
    - lack of equipment and infrastructure.

C. Weak technical guidance, along with strengthening of other elements of the service delivery level, or other levels at which constraints to improving access to health care lie:

- community/household level;
- health sector policy and strategic management level;
- public policies cutting across sectors; and
- environmental and contextual characteristics.

We addressed the second objective of synthesising the effectiveness of included interventions by using narrative synthesis to outline the direction of the effect, size of the effect and whether the effect was consistent across studies within and between the three-tier classification described above and more fully in the results section. We summarised the data using tables.

In addition to classifying the interventions according to the targeted constraint and level of the health system it aimed to improve, we also used narrative synthesis to systematically describe and compare interventions based on their characteristics (Thomas et al., 2012). Characteristics are shown in Appendix 5 and included:

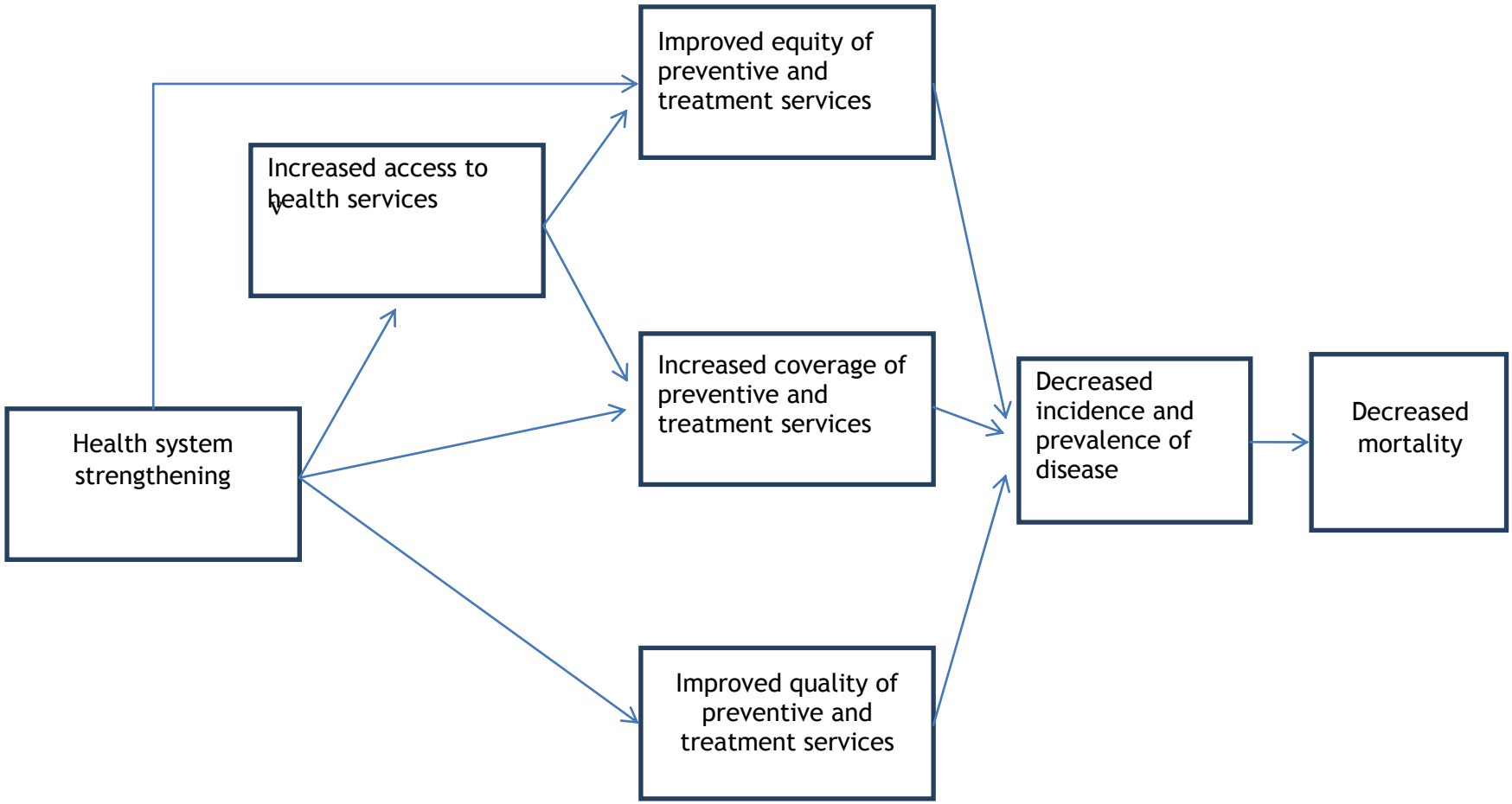
- participants;
- intervention:
    - implementers/involvement of intervention site staff;
    - location of implementation (on-site, external);
    - length of implementation (one-off, multiple, number of days);
    - delivery method of intervention (e.g. face-to-face, remote);
    - components of the intervention;
- setting (urban/rural) and country/geographical region;
- disease area.

2.4.3 Mapping the epidemiological evidence to a logic model of pathways of influence

In this review, we also aimed to map out the availability of robust epidemiological evidence for the effectiveness of interventions on outcomes, by means of a logic model showing anticipated pathways of influence. The pathways through which health service delivery improvement is anticipated to affect outcomes, in particular survival impact, are complex. However, a simple starting point is shown in Figure 2.2.

This shows that the inputs and processes of strengthening the health system are anticipated to improve outputs, and in turn influence outcomes of access, coverage, quality and equity, and eventually have an impact on reduced incidence (prevention) and prevalence (treatment) of disease, which in turn have an impact of reduced mortality. Evidence from the included studies was used to suggest a more comprehensive model of pathways of influence in section 5.3 of the results (Figure 2.2).

Figure 2.2: Simple framework for causal link between health system strengthening and mortality



3. Results

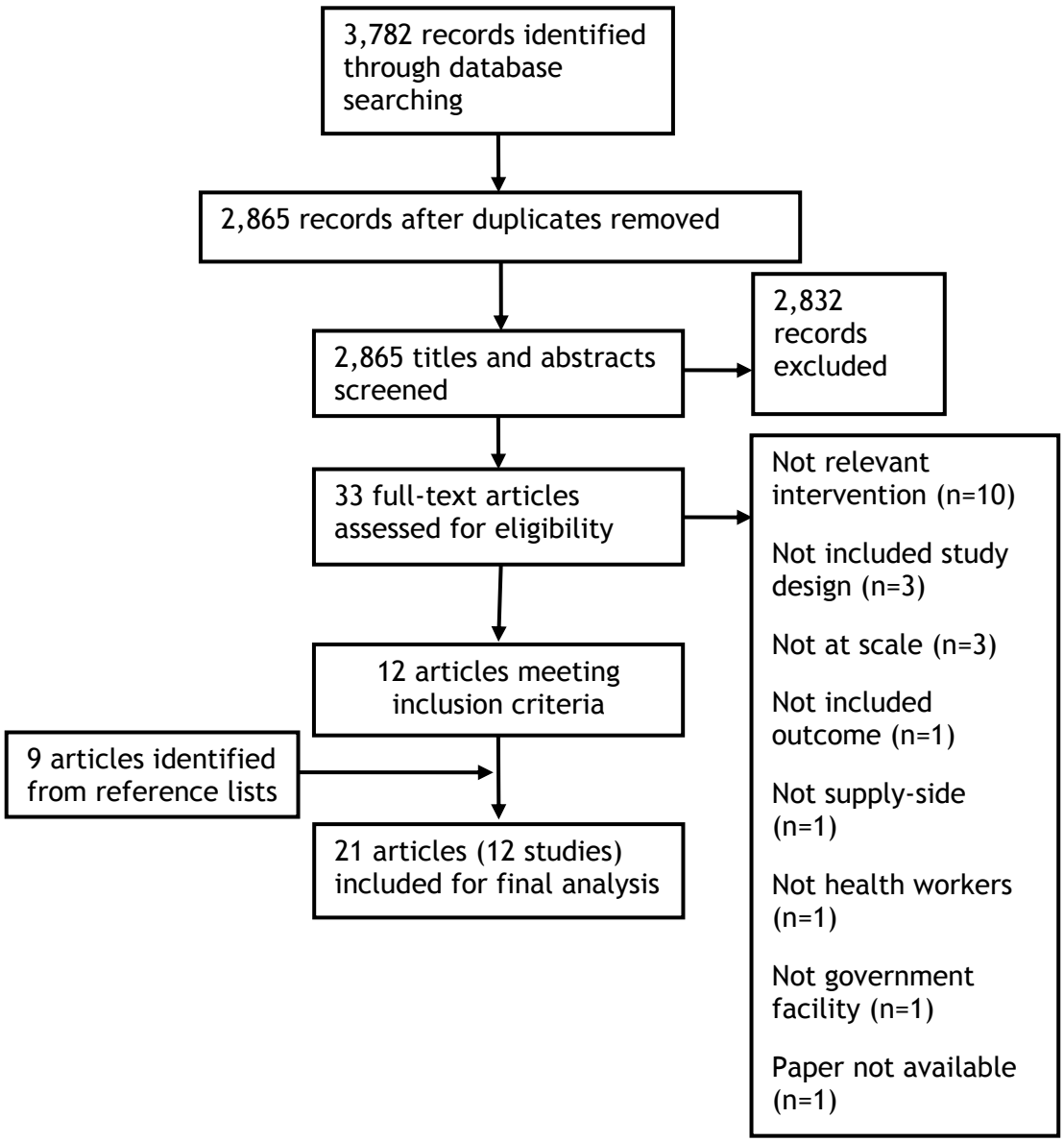
3.1 Database search results and screening

The search results and screening process are shown in the PRISMA flow chart (Figure 3.1).

In total, 2,865 papers were identified. Two reviewers (BW and LSP) independently screened the titles and abstracts of these records against pre-specified eligibility criteria, and 33 papers were retained for full paper screening, after which a further 21 were excluded (Figure 3.1 and Appendix 6).

In total, 12 papers were retained, and a further nine were identified from reference lists of included papers, giving 21 papers for full analysis. Of these, 15 were index papers reporting main study outcomes and six were satellite papers, which were included to complement information reported in the index papers. The final 21 papers covered a total of 12 studies.

Figure 3.1: PRISMA flow chart of literature search and screening



3.2 Quality of included studies

- Ten studies used a cluster randomised trial (CRT) design, and two a non-randomised quasi-experimental cluster design (both, as expected, were categorised as having a high risk of bias)
- Of the ten CRTs, only one was at low risk of bias across all categories, with an additional four at low risk of bias across the majority of categories

The papers were assessed under the categories of protection against contamination, randomisation (allocation concealment and random sequence generation), incomplete outcome assessment, blinding of outcome measurement, and similarity of groups at baseline, and they were categorised as done, not done, or unclear (when the details were not reported in the paper), corresponding to low, high or unknown risk of bias (Higgins and Green, 2011). In addition we recorded whether the study reported an appropriate sample size calculation - in particular whether adjustment for clustering in the calculation was carried out.

Table 3.1 provides an overview using a traffic light colour-coding system (green for 'done', red for 'not done' and amber for 'unclear') and Table 3.2 shows full details. In summary:

- Most studies reported follow-up or presented trial profiles to allow assessment of whether there was incomplete outcome assessment.
- Where this was available, follow-up was generally above 80 percent. However, some studies did not present this information and are therefore graded 'not clear'.
- On the whole, blinded assessment of the main study outcome was not carried out in these studies. However, as this outcome was frequently quality of care and assessed using structured survey tools or checklists, outcome assessment is unlikely to have been subjective, reducing potential measurement bias.
- Similarity between groups at baseline was reported with disappointingly low frequency, making it hard to assess success of randomisation that may be a particular issue in cluster designs.

Table 3.1: Overview of quality of included studies

Reference	Contamination	Randomisation	Incomplete outcome data	Blinding of outcome assessment	Similarity at baseline	Power
Arifeen et al. (2004a, b)	✓	✓	✓	✗	✓	✓
Armstrong Schellenberg et al. (2004a, b)	✓	✗	✓	✗	✗	✗
Ayieko et al. (2011)	✓	✓	✓	✗	✓	✓
Bradley and Igras (2005)	✓	✗	✓	✗	?	✗
Gilroy et al. (2004)	✓	✓	?	?	✗	✓
Gulmezoglu et al. (2007)	✓	✓	✓	✗	✓	✓
Jennings et al. (2010)	✓	✓	✓	✗	✗	✓
Ngasala et al. (2008)	✓	✓	✓	✗	✗	✓
Rowe et al. 2009)	✓	✓	✗	✗	?	?
Winch et al. (2003)	✓	✓	?	✗	✗	✓
Zaman et al. (2008)	✓	✓	✓	✓	✓	✓
Zurovac et al. (2011)	✓	✓	✓	✗	✓	✓

Table 3.2: Study quality and risk of bias, based on Cochrane EPOC checklist

Reference and study design	Adequate protection against contamination	Randomisation: allocation concealment and random sequence generation	Incomplete outcome data (follow-up)	Blinding of primary outcome assessment	Similarity of outcome and other characteristics at baseline between intervention and comparison groups	Power calculation
Arifeen et al. (2004a, b) CRT (n=20 catchment areas)	Done - low risk of bias	Done - low risk of bias	Done- low risk of bias - 100% follow-up, trial profile presented	Not done. Quality of case management based on standardised data collection form. Mortality assessed by verbal autopsy in household census	Done - low risk of bias. No differences between areas at baseline, except in use of sanitary toilets (higher in IMCI areas)	Done
Ayieko et al. (2011) CRT (8 district hospitals)	Done - low risk of bias	Done - low risk of bias	Done- low risk of bias - 100% follow-up, trial profile presented	Not done, but quality of care collected from patient records using abstraction form	Done - low risk of bias	Done
Gilroy et al. (2004) CRT (n=10 community health centres)	Done - low risk of bias	Done - low risk of bias	Not clear- not specified in the paper and no trial profile presented	Not clear- not specified in the paper. Data collected using standardised 17-item checklist	Not done - high risk of bias. Data presented, but important differences in socio-economic status between groups at baseline	Done

Reference and study design	Adequate protection against contamination	Randomisation: allocation concealment and random sequence generation	Incomplete outcome data (follow-up)	Blinding of primary outcome assessment	Similarity of outcome and other characteristics at baseline between intervention and comparison groups	Power calculation
Gulmezoglu et al. (2007) CRT (n=18 hospitals Thailand; 22 hospitals Mexico)	Done - low risk of bias	Done - low risk of bias	Done- low risk of bias - all included hospitals contributed data	Not done, but collected by field workers not involved in implementation	Done - low risk of bias	Done
Jennings et al. (2010) CRT (14 maternities)	Done - low risk of bias	Done - low risk of bias	Done- low risk of bias - all included health centres contributed data	Not done, but patient consultation observations based on a standardised checklist	Not done - high risk of bias. Data presented but important differences in characteristics of pregnant women in intervention and comparison centres, in education and timing of presentation for ANC	Done

Reference and study design	Adequate protection against contamination	Randomisation: allocation concealment and random sequence generation	Incomplete outcome data (follow-up)	Blinding of primary outcome assessment	Similarity of outcome and other characteristics at baseline between intervention and comparison groups	Power calculation
Ngasala et al. (2008) CRT (n=16 health facilities)	Done - low risk of bias	Done - low risk of bias	Done- low risk of bias - >99% follow-up	Done for malaria microscopy - health facility reading unblinded but study re-reading (definitive) blinded. Not done for quality of care	Not done - high risk of bias. Data presented but important differences in prevalence of febrile children between groups – there was lower prevalence in the group trained in the clinical algorithm only	Done
Rowe et al. (2009) CRT (n=8 districts)	Done - low risk of bias	Done - low risk of bias	Not done- high risk of bias - data collected from between 55% and 81% of eligible health facilities, drop off due to health facilities without >1 ill child consultation available	Not done- high risk of bias, but patient consultation observations based on a standardised checklist	Not clear, data not presented	Not clear - not presented in paper

Reference and study design	Adequate protection against contamination	Randomisation: allocation concealment and random sequence generation	Incomplete outcome data (follow-up)	Blinding of primary outcome assessment	Similarity of outcome and other characteristics at baseline between intervention and comparison groups	Power calculation
Winch et al. (2003) CRT (n=10 health zones)	Done - low risk of bias	Done - low risk of bias	Not clear- not specified in the paper and no trial profile presented	Not done, but data collected using structured questionnaires during follow-up day 5 with care givers	Not done - high risk of bias. Data presented but important differences in characteristics of caregiver's education in intervention and comparison areas	Done
Zaman et al. (2008) CRT (n=40 health facilities)	Done - low risk of bias	Done - low risk of bias	Done- low risk of bias - >80% followed up	Done- low risk of bias	Done - low risk of bias	Done
Zurovac et al. (2011) CRT (n=11 districts)	Done - low risk of bias	Done - low risk of bias	Done- low risk of bias - 80% followed up immediately, although 75% of enrolled followed up 6 months after intervention.	Not done. Data collected during structured interviews with care givers who were blinded to the intervention	Done - low risk of bias	Done

Reference and study design	Adequate protection against contamination	Randomisation: allocation concealment and random sequence generation	Incomplete outcome data (follow-up)	Blinding of primary outcome assessment	Similarity of outcome and other characteristics at baseline between intervention and comparison groups	Power calculation
Non-randomised designs						
Armstrong Schellenberg et al. (2004a, b) Quasi-experimental cluster design (n=4 districts)	Done - low risk of bias	Not done (IMCI already implemented in certain districts of Tanzania at time of study initiation)	Done- low risk of bias - consultations observed in 97% of health facilities included	Not done, but patient consultation observations based on a standardised data collection form	Not done. IMCI implemented in comparison districts two years after study (2002); implemented in intervention districts from 1997	Not done, all facilities in both comparison districts (n=16 and 19), and random sample of 20 facilities from both interventio n districts
Bradley and Igras (2005) Non-randomised cluster design (n=4 districts Kenya; 8 districts Guinea)	Done - low risk of bias	Not Done - high risk of bias (non-randomised design). Authors do not provide rationale for not using randomisation	Done- low risk of bias - all included health centres contributed data	Not done, but patient consultation observations based on a standardised checklist	Not clear, data not presented, but it was reported in text that data were similar at baseline (however, reference unavailable online)	Not done

Notes: Protection against contamination of comparison groups with the intervention was not presented as all studies scored 'done' through the use of cluster designs. Selective outcome reporting not apparent from papers, but protocols of studies not obtained for comparison. Rowe et al. reported per protocol rather than intention to treat analysis due to contamination of comparison areas with IMCI due to slow roll-out of intervention.

3.3 Overview of included studies

- The twelve studies were conducted in nine countries in West Africa, East Africa, Asia, Latin America
- Eight studies were in more than one district
- Eight included health workers managing children at first- or referral-level facilities
- All aimed to improve the diagnosis and management of infectious disease in children
- All used a cluster design; 10 were cluster randomised controlled trials

Details of the included studies are shown in Table 3.3.

3.3.1 Geographical setting and scale of studies

The 12 included studies were conducted in nine countries - six low income, two lower middle income countries, and one upper middle income country (part of a multi-country study with a lower middle income country).

The majority of studies were conducted in rural settings (7/12), one in an urban setting, and four in mixed urban/rural settings.

Four studies were conducted in one district only, where a district is defined as the lowest administrative level. In these four studies, comparisons were at the sub-district level.

Six further studies were conducted in more than one district (range 2-11 districts), and two studies were conducted in more than one sub-national region (range 2-4 regions).

3.3.2 Characteristics of health facilities and health care workers

This review was concerned with interventions to strengthen national health service delivery by state-employed front-line health workers, and thus, all studies included state-provided health care facilities.

Primary-level facilities (health centres and dispensaries) were the most common type of facility included (8/14).[3] Maternity units were included in two studies (Gulmezoglu et al., 2007, Jennings et al. 2010). District hospitals were included in one study introducing new guidelines in Kenya (Ayieko et al., 2011). One study from Mali improving treatment for malaria involved community-based care in the form of village drug kits (Winch et al., 2003). The majority of studies included health workers managing children at first- or referral-level facilities, such as doctors, clinical officers, nurses and in some cases nursing aides.

3.3.3 Disease areas targeted

All included studies focused on paediatric care, with some newborn, maternal and antenatal care. In total, eight included a malaria focus, five included an acute respiratory illness (ARI) focus, and five a diarrhoea focus; four focused on reproductive, maternal and neonatal health; one study each focused on child nutrition and uptake of HIV testing.

[3] Note: denominator is 14 rather than 12, as the two multi-country studies included more than one type of state-provided health care facility.

3.3.4 Study design

All studies were carried out using a cluster design. Ten used a cluster randomised controlled trial (CRT) design. One study used a quasi-experimental study design and one a non-randomised cluster trial study design (Armstrong Schellenberg et al., Armstrong Schellenberg et al. 2004a, b, Bradley and Igras, 2005 respectively).

3.3.5 Funders and affiliations of research teams

Appendix 8 shows details of the funders of included studies and the affiliations of the research teams, based on details provided in the publications.

Some studies were funded by more than one organisation, and in total 17 funders of 12 studies were identified. These were principally: bilateral donors (USAID, Swedish international Development Cooperation Agency (SIDA/SAREC) and The International Development Research Centre, Canada) contributing to 11 studies; charitable foundations (Wellcome Trust UK and The Bill and Melinda Gates Foundation) contributing to two studies each; and Multilateral donors (UNDP, UNFPA, WHO, World Bank Research Training in Human Reproduction and UNICEF) contributed to one study each.

Based on authors' affiliations, the research teams were principally based in American, Swedish, British or Canadian academic research institutions, however authors from academic research institutions in the countries of study were also common. Some authors were included from bi- or multi-lateral donor organisations and some charities, as well as a US government agency (Centers for Disease Control). Authors from government positions within countries of study were not common.

3.4 Objective 1: Characterising the interventions

- A wide variety of interventions were documented from the 12 included studies
- A three-tier categorisation of interventions was created, based on the WHO building blocks framework and the framework by Hanson and colleagues
- Three categories were identified and described, lying along a spectrum of increasing incorporation of other elements of the health service delivery level, and/or other levels at which constraints to improving access to health care act

Given the perspective of this review, we have used a classification based on the extent to which an intervention addressed improving the service provided by front-line workers during their interaction with users, located within the context of levels of the health system more generally, as defined by Hanson and colleagues (Hanson et al., 2003).

To clarify, we present our visual interpretation of Hanson and colleagues' framework in Figure 3.2. This shows the health system comprised of:

- The 'health services delivery' and 'health sector policy and strategic management' levels (levels II and III)
- These are embedded within 'public policies cutting across sectors' (level IV)
- This in turn is located within the 'wider context' (level V).

- The 'community and household' level (level I) is shown by a black-outlined box, cutting across other levels.

Figure 3.2 additionally incorporates the six WHO building blocks of a health system:

1. Service delivery
2. Workforce (human resources)
3. Information
4. Supplies (medicines and technologies)
5. Financing
6. Leadership and governance

We use a three-tier classification and present the results of the studies under the same classification together (Table 3.3 and Figure 3.2):

A. Interventions that address improving the service provided by front-line workers during their interaction with users through technical guidance, without wider health system involvement

B. Interventions that address improving the service provided by front-line workers during their interaction with users through technical guidance

   AND

   through strengthening of other elements of the service delivery level:

   - shortage, distribution and training of staff;
   - programme management and supervision;
   - drug and medical supplies;
   - equipment and infrastructure.

C. Interventions that address the improving the service provided by front-line workers during their interaction with users through technical guidance

   AND

   through strengthening of other elements of the service delivery level, or other levels at which constraints to improving access to health care lie:

   - community/household level;
   - health sector policy and strategic management level;
   - public policies cutting across sectors;
   - environmental and contextual characteristics.

Table 3.3: Overview of included studies

Reference	Country and setting	Intervention overview	Characteristics of intervention setting and participants	Scale (n)
(A) Improving the service provided by front-line workers during their interaction with users				
Gilroy et al. (2004)	Mali; rural	IMCI: *Health service delivery level (II) Training* Implementation of the case management improvement element of IMCI. 11-day WHO IMCI case management improvement course of interactive sessions, including role play for practising communication skills, and one supervisory visit about 1 month after completion of training.	Government community health centres; head nurses	District (n=1)
Gulmezoglu et al. (2007)	Thailand and Mexico; peri-urban Thailand; urban Mexico	WHO Reproductive Health Library: *Health service delivery level (II) Training* 3 seminar sessions over 6 months on accessing and using the WHO's RH Library electronic resources; computers and some printed materials.	Government (Thailand), Social Security (Mexico) maternity hospitals; doctors, midwives, medical students	District (n=1, both)
Ngasala et al. (2008)	Tanzania; rural	Malaria treatment and diagnosis: *Health service delivery level (II) Training* Training in malaria management (5 days), or additionally trained in malaria microscopy diagnosis (5 days) and provided with a microscope and slide supplies.	Government first-level facilities; HW managing children, laboratory staff	District (n=2)
Zaman et al. (2008)	Pakistan; urban	IMCI nutrition: *Health service delivery level (II) Training* IMCI nutrition counselling module training 5 half-day sessions over 5 days.	Government health centres; lady health visitors	District (n=1)

Reference	Country and setting	Intervention overview	Characteristics of intervention setting and participants	Scale (n)
		Training included practical sessions; a job aid of the IMCI nutrition counselling card in the local language was also provided.		
Zurovac et al. (2011)	Kenya; rural	SMS malaria: *Health service delivery level (II) Training* Text message reminders twice a day, 5 days a week, for 6 months. Messages contained reminders for the management of malaria and a motivational quote.	Government first-level facilities (health centres and dispensaries); health workers managing children	District (n=11)
(B) Improving the service provided by front-line workers during their interaction with users alongside strengthening at the service delivery level				
Ayieko et al. (2011)	Kenya; rural	Implementation of guidelines: *Health service delivery level (II) Training* Staff were trained in emergency triage, assessment, treatment and admission care (ETAT+) guidelines during 17 lectures and practicals over 5.5 days and provided with job aids (paediatric admission record forms and reference text books). *Health service delivery level (II) Management and Supervision* External enhanced supportive supervision was planned every 2-3 months, and a full-time in-facility facilitator was nominated to encourage use of guidelines and troubleshoot on-site problems, supported by external telephone supervision every 1-2 weeks. Additionally, six-monthly health facility surveys were carried out, with face-to-face feedback provided over 18 months.	Government district hospitals; HW managing children	Region (n=4)

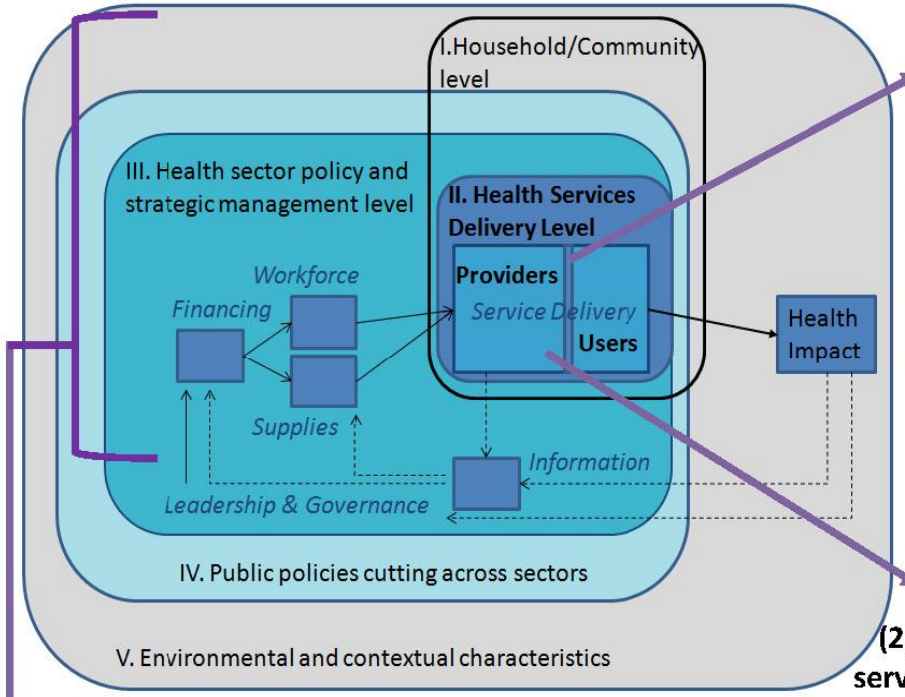
Reference	Country and setting	Intervention overview	Characteristics of intervention setting and participants	Scale (n)
Bradley and Igras (2005)	Guinea and Kenya; mixed urban and rural	Health centre quality improvement: *Health service delivery level (II) Training* Staff underwent training on supervision, communication, counselling, immunisation and infection prevention. *Health service delivery level (II) Management* Staff participated in sequential self-administered quality improvement assessment sessions over 2-3 days, followed by planning sessions for an action plan with named responsible individuals and time frames. The process was repeated every 3-4 months over 15 months.	Government health centres; all staff including non-professional and district managers	District (n=8 Guinea; n=4 Kenya)
Jennings et al. (2010)	Benin; rural	ANC job aids: *Health service delivery level (II) Training* Training for 3 days in the use of job aids to guide counselling during antenatal visits. Training was interactive and a supervision visit was included 1 month after. *Health service delivery level (II) Management* Maternity units were also encouraged to form teams to identify organisational changes that were needed in order to successfully implement use of job aids	Government maternity units; nurses and midwives	Region (n=2)
Rowe et al. (2009)	Benin; mixed urban and rural	Post IMCI supervision: *Health service delivery level (II) Training* 11 day IMCI case management improvement course.	Government and licensed private first-level facilities; HW managing children (excluding nursing aides)	District (n=8)

Reference	Country and setting	Intervention overview	Characteristics of intervention setting and participants	Scale (n)
		Health service delivery level (II) Supervision Enhanced supportive supervision of 2 supervision visits every 3 months planned. Job aids in the form of a counselling guide and patient register forms were provided. Supervision of supervisors was also included, and face-to-face feedback to facilities by supervisors was incorporated as the implementation progressed.		
(C) Improving the service provided by front-line workers during their interaction with users and strengthening at the service delivery, community/household, or health sector policy level				
Arifeen et al. (2004a, b)	Bangladesh; rural	IMCI: *Health service delivery level (II) Training* All 3 elements of IMCI were implemented: training 11-day IMCI case management improvement course, job aids (counselling cards and referral forms) and supplies (scales, thermometers and timers to measure respiration rates) *Health service delivery level (II) Supplies* Provision of essential drugs and supplies and introduction of facility supply tracking and reporting system *Health service delivery level (II) Management and Referral* Referral forms were linked to an introduced referral system, with sensitisation activities for referral facility staff *Health service delivery level (II) Shortage and Distribution of Staff* Guidelines for the management of pneumonia at first-level facilities were	Government first-level facilities; HW managing children	District (n=1)

Reference	Country and setting	Intervention overview	Characteristics of intervention setting and participants	Scale (n)
		amended in intervention areas to allow treatment prior to referral of severe cases, and later a new cadre of community health worker, with associated training and supervision, was introduced to manage uncomplicated pneumonia at the community level *Community/Household level (I)* Community mobilisation included training of local imams and the use of community theatre to encourage health seeking		
Armstrong Schellenberg et al. (2004)	Tanzania; rural	IMCI: *Health service delivery level (II) Training* 2 of the 3 IMCI elements implemented with 11-day IMCI case management improvement course, including 30% of time spent on clinical practice, followed by a supervision visit within a month. *Health sector policy and strategic management level (III)* Increased autonomy for district council health management teams on budgeting and planning, linked to availability of district level burden of disease data, a budget modelling tool, and simulated basket funding – all linked to the Tanzania Essential Health Intervention Project.	Government first level and referral facilities; HW managing children	District (n=4)
(Winch et al., 2003)	Mali; mixed urban and rural	Community treatment of malaria: *Health service delivery level (II) Training* Training in age-specific dosing of chloroquine over 5 days. Job aids included visual aids and danger signs requiring referral.	Government community drug kits; community drug kit manager	District (n=1)

Reference	Country and setting	Intervention overview	Characteristics of intervention setting and participants	Scale (n)
		Health service delivery level (II) Management and Referral Raised awareness with head nurses of intervention area community health centres (see study above by Gilroy et al., 2004) which would receive referred children. *Community/Household level (I)* Included village meetings to encourage use of the village drug kits		

Figure 3.2: Authors' visual representation of the levels of constraints to improving access to priority health care (Hanson et al., 2003) and the WHO health system building blocks (World Health Organization, 2007)



(1) Health service delivery- improving service provided by front-line workers during their interaction with users by technical guidance only

Text message reminders for malaria management (Kenya)

Access to WHO Reproductive Health Library resources (Mexico and Thailand)

Clinical algorithm and microscopy training for malaria management (Tanzania)

IMCI improving case management (Mali)

IMCI Nutrition counselling (Pakistan)

(2) Health service delivery- Improving the service provided by front-line workers during their interaction with users and strengthening at the service delivery level

Job aids to improve ANC counselling (Benin)

Implementing on-going post- IMCI enhanced supervision (Benin)

Implementing pediatric and newborn care guidelines (Kenya)

Quality improvement intervention (Guinea and Kenya)

(3) Household/ community & Health service delivery & Health sector policy & strategic management level

Training for community malaria treatment (Mali)

IMCI improving case management; strengthening health system (Tanzania)

IMCI improving case management; strengthening health system; community mobilisation (Bangladesh)

(A) Improving the service provided by front-line workers during their interaction with users

Five studies involved interventions to improve the delivery of services by front-line health workers without wider health service delivery components.

In Kenya, health workers responsible for managing paediatric outpatient cases at first-level facilities in both intervention and comparison areas received training in malaria case management and dissemination of national guidelines. Health workers in the intervention areas were then sent text message reminders twice a day, five days a week, for six months. The messages contained reminders for the management of malaria that were based on national guidelines, as well as a motivational quote (Zurovac et al., 2011).

In a multi-country study from Mexico and Thailand, doctors, midwives, interns and medical students from maternity hospitals were offered three on-site seminar sessions over six months on accessing and using the WHO's Reproductive Health Library electronic resources; these contained up-to-date evidence-based recommendations for obstetric and neonatal care relevant to developing countries. They also received computers and some printed materials (Gulmezoglu et al., 2007).

A study from Tanzania, where all staff from district first-level and hospital facilities involved in paediatric care and laboratory assistants were either (i) trained on-site in clinical malaria management only (symptom recognition, treatment, recognition of danger signs, importance of referral) for five days, or (ii) additionally trained in malaria microscopy diagnosis (slide preparation, slide reading, microscope maintenance) for five days and provided with a microscope and slide supplies (Ngasala et al., 2008).

A Malian study included the implementation of the case management improvement element of IMCI. Head nurses of community health centres attended the standard 11-day WHO IMCI case management improvement course of interactive sessions, including role play for practising communication skills, held off-site. About one month after completion of training, a supervisory visit was carried out in health centres by training leaders (Gilroy et al., 2004).

Zaman and colleagues (2008) described the use of the IMCI nutrition counselling module to train lady health visitors from health centres in Pakistan in appropriate infant feeding and counselling skills during five half-day sessions over five days. The location of the training was not specified. The training included practical sessions, and a job aid of the IMCI nutrition counselling card in the local language was also provided.

(B) Improving the service provided by front-line workers during their interaction with users, with additional strengthening at the service delivery level

Four studies were classified as aiming to improve the delivery of services by front-line workers to users, along with strengthening of one other element at the health service delivery level.

In Benin, nurse midwives from maternity units were trained over three days in the use of job aids to guide counselling during antenatal visits. The location of the training was not specified, but sessions were interactive, with both individual and group role play with feedback. A supervision visit was undertaken one month after training, and included discussion of difficulties encountered during implementation. Job aids were grouped by stage of pregnancy. Maternity units

were also encouraged to form teams to identify organisational changes that were needed in order to successfully implement the use of the job aids (Jennings et al., 2010).

Rowe and colleagues (2009) reported on a study from Benin, where ongoing enhanced supportive supervision for health workers managing paediatric patients was implemented in first-level government and licensed private health facilities, following IMCI training. A protocol for enhanced supervision, including a supervision session checklist, was developed in collaboration with intervention sites and district officials, and two on-site supervision visits every three months were planned. Job aids in the form of a counselling guide and patient register forms were provided. Supervision of supervisors was also included, through the appointment of a senior paediatrician to undertake this, and face-to-face on-site feedback to facilities by supervisors was incorporated as the implementation progressed.

In Kenya, district hospital staff providing paediatric and neonatal care, along with district managers, were involved in comparing two intensities of intervention to introduce emergency triage, assessment, treatment and admission care (ETAT+) guidelines. Staff in intervention hospitals were trained on-site during 17 lectures and practicals over 5.5 days, and provided with a variety of job aids, including paediatric admission record forms and reference textbooks. External enhanced supportive supervision was planned for facility staff every two to three months, and a full-time in-facility facilitator was nominated to encourage the use of the guidelines and to troubleshoot on-site problems, supported by external telephone supervision every one to two weeks. Additionally, six-monthly health facility surveys on resource availability and challenges to implementing guidelines were carried out, with face-to-face feedback provided in facilities to staff and district managers, over a period of 18 months. Comparison hospitals received written feedback on the six-monthly surveys, and were provided with the guidelines and job aids but did not receive any further support (Ayieko et al., 2011).

Bradley and Igras (2005) described a multi-country study from Guinea and Kenya where all health centre staff, including non-professional staff, and district managers participated in sequential self-administered quality improvement assessment sessions carried out on-site over two to three days, followed by action planning sessions complete with named responsible individuals and time frames. The process was repeated every three to four months over a 15-month period. Staff also underwent some training on supervision, communication, counselling, immunisation and infection prevention.

(C) Improving the service provided by front-line workers during their interaction with users, in addition to strengthening at the service delivery and/or community/household or health sector policy levels

Three studies were classified as aiming to improve the delivery of services by front-line workers to users, along with strengthening at the health service delivery level, and/or strengthening at the household or community, or health sector policy level (Hanson et al., 2003).

In Mali, community-based lay health workers (village drug kit managers) were trained over five days in age-specific dosing of chloroquine (CQ) for malaria treatment, counselling caregivers on the administration of CQ in the home, and conditions for referral to the health centre. A referral mechanism from drug kit managers to the first-level health facilities was established. Job aids included visual aids for CQ treatment and danger signs requiring referral. This study also

raised awareness with head nurses of intervention area community health centres (as previously described by Gilroy et al., 2004) which would receive referred children, and included village meetings to encourage use of the village drug kits (Winch et al., 2003).

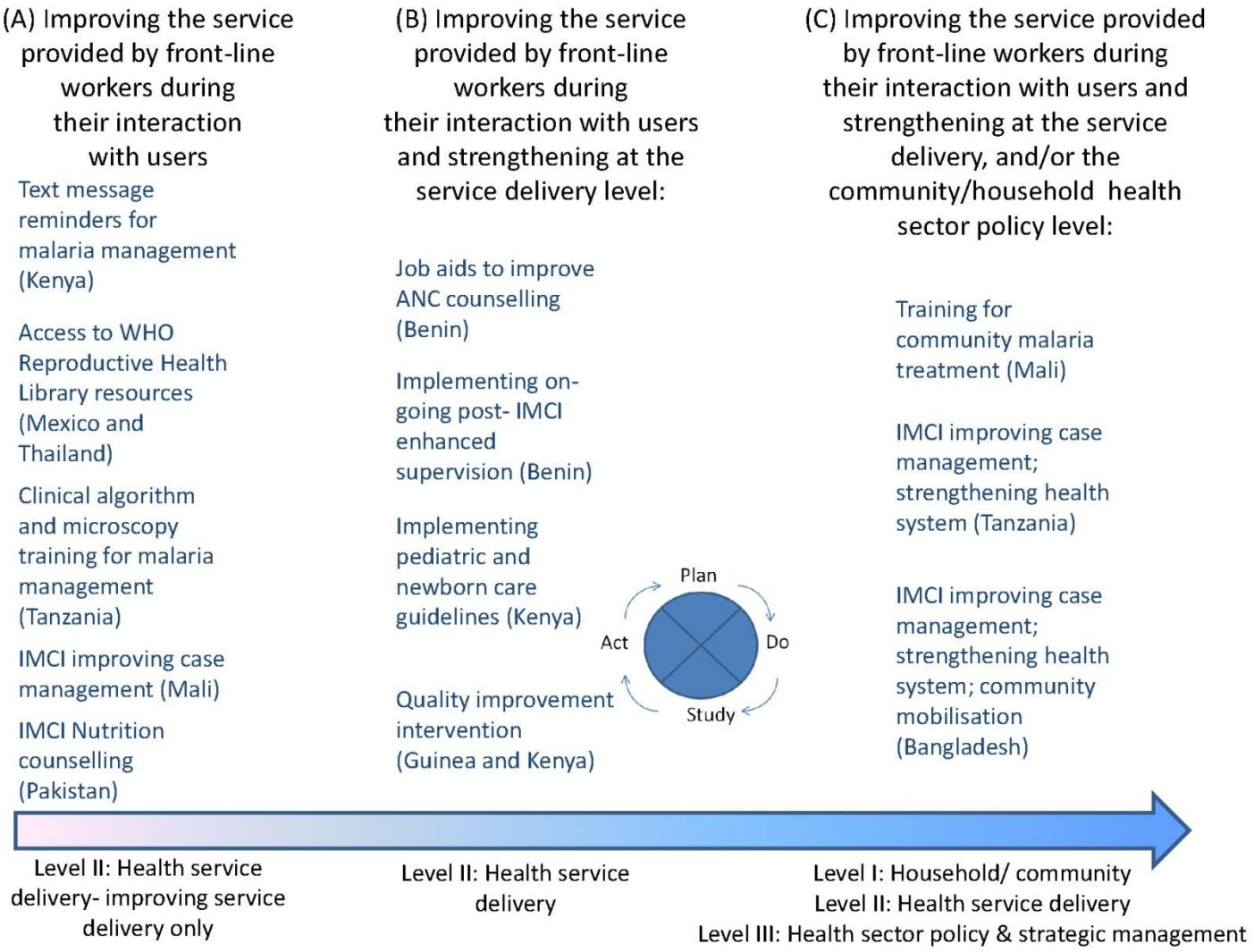
In Tanzania, two of the three IMCI elements were implemented in health centres and dispensaries. Staff managing paediatric cases were trained in the WHO's 11-day IMCI case management improvement course, including 30 percent of time spent on clinical practice, followed by a supervision visit by the trainer within a month of training. Monthly post-training supervision visits included observations of case management with immediate feedback and supportive interaction, including discussion of barriers to full implementation. Health system support at the health sector policy and strategic management level was linked to the Tanzania Essential Health Interventions Project (TEHIP), which included increased autonomy for intervention district council health management teams on budgeting and planning, through district-level burden of disease information, a budget modelling tool, and simulated basket funding (Armstrong Schellenberg et al., 2004a, b).

Implementation of all three elements of IMCI in Bangladesh included training all staff managing paediatric cases at first-level facilities in the WHO's 11-day IMCI case management improvement course, the provision of a variety of job aids, such as counselling cards and referral forms, as well as supplies of scales, thermometers and timers to measure respiration rates. Monthly post-training supervision visits included observations of case management, with immediate feedback and supportive interaction, including discussion of barriers to full implementation. Wider health system support interventions included provision of essential drugs and supplies, as well as the introduction of a facility supply tracking and reporting system. Referral forms were linked to a new referral system, with sensitisation activities for referral facility staff. Guidelines for the management of pneumonia at first-level facilities were amended in intervention areas to allow treatment prior to referral of severe cases, and later a new cadre of community health worker, with associated training and supervision, was introduced to manage uncomplicated pneumonia at the community level. Community mobilisation included training of local imams and the use of community theatre to encourage health seeking (Arifeen et al., 2004a, b).

Figure 3.3 illustrates how the interventions included in this review can be conceptualised along a spectrum of increasing incorporation of other elements of the health service delivery level, and/or other levels at which constraints to improving access to health care act (Hanson et al., 2003).

The interventions range from reminders and refresher training through training on new health-specific technical content, to complex multi-faceted interventions that incorporate a plan-do-study-act approach of cycles of implementation, assessment and feedback to improve service delivery.

Figure 3.3: Spectrum of interventions to improve national health service delivery



3.5 Objective 2: Outcomes reported: coverage, equity, quality and impact

- Five studies reported on nine coverage outcomes, including health seeking, breastfeeding and management of childhood diseases
- None of the 12 studies reported outcomes relating to access
- The majority of studies (11/12) reported quality outcomes
- One study reported outcomes relating to equity
- Three studies reported on two outcomes relating clinical impact
    - Anthropometric outcomes (3 studies)
    - Under-five mortality (2 studies)

Details of the included studies including the results, outcomes reported, measurement of outcomes, sample sizes and study designs are shown in Tables 3.4-3.6.

3.5.1 Coverage

Five studies reported on nine coverage outcomes, including health seeking, breastfeeding and management of childhood diseases.

- In their respective IMCI evaluations, Armstrong Schellenberg and colleagues (2004a, b), as well as Arifeen and colleagues (2004a, b) reported coverage of **health seeking behaviour** for illness in under-fives during the past two weeks, based on interviews with care takers during household surveys.
- These two studies also reported the prevalence of **exclusive breastfeeding** among infants aged less than four months (Tanzania) and six months (Bangladesh).
- These two IMCI evaluations also reported the prevalence of correct treatment of priority childhood illnesses, including **malaria, pneumonia and diarrhoea** (Bangladesh only), based on observations of case management.
- Ngasala and colleagues (2008) reported coverage of **antimalarial prescriptions** among febrile under-fives.
- Winch and colleagues (2003) reported coverage of **home treatment with chloroquine** among febrile under-fives.
- Ayieko and colleagues (2011) reported coverage of provider-initiated **HIV testing** and counselling, **Vitamin A** administration on admission to the hospital, and documentation of a child's **immunisation status**, based on hospital records.

3.5.2 Access

None of the 12 studies reported outcomes relating to access.

3.5.3 Quality

The majority (11/12) reported quality outcomes, including:

- five studies reporting quality of case management of childhood illness;
- three reporting quality of malaria treatment in children;
- one reporting quality of nutrition counselling in children;
- one reporting quality of antenatal counselling in pregnant women; and
- one reporting quality of obstetric practices in pregnant women.

These quality outcomes were principally measured through direct observation (7/11); however, in two studies, this was measured by referring to routine records (Ayieko et al., 2011, Gulmezoglu et al., 2007), and in two by caregiver recall (Winch et al., 2003, Zurovac et al., 2011). Many created scores or indices, but methods were not standardised across studies.

3.5.4 Equity

Only the Tanzanian evaluation of IMCI reported outcomes of equity. Masanja and colleagues reported that socioeconomic quintiles of household were developed using principal component analysis, and a concentration index was calculated. The difference of difference in the concentration index of ten outcomes was calculated (Masanja et al., 2005, a satellite paper relating to Armstrong Schellenberg et al, 2004a, b).0

3.5.5 Impact

Three studies reported on two outcomes of clinical impact:

1. Anthropometric outcomes: height-for-age, weight-for-age, and weight-for-height assessed during household surveys.
2. Under-five mortality reported from IMCI evaluations:
    - In Bangladesh, Arifeen and colleagues (2004a, b) collected mortality based on verbal autopsy, and reported under-five mortality rate excluding neonatal mortality (i.e., 7 days- 5 years), per 1,000 children.
    - In Tanzania, Armstrong Schellenberg and colleagues (2004a, b) reported under-five mortality rates based on demographic surveillance of district-specific mortality rates between mid-2000 and mid-2002, five years after the intervention.

3.6 Objective 2: Effectiveness of interventions to strengthen health service delivery

- It was not appropriate to pool effect estimates in a meta-analysis
- The effectiveness of interventions on coverage, quality, equity and impact are discussed by each of the three classifications of intervention
- On the whole, studies classified into categories B and C showed more consistently improved outcomes than studies whose intervention consisted of technical guidance alone (category A). This was particularly the case for quality outcomes

The effectiveness of interventions on coverage, quality, equity and impact are discussed in relation to each of the three classifications of intervention (see

Section 3.4 above). The effectiveness of interventions on outcomes of access is not discussed, as none of the included studies reported outcomes of access.

(A) Improving the service provided by front-line workers during their interaction with users

Details of studies in this category are shown in Table 3.4.

Effect on coverage

Ngasala and colleagues (2008) reported a CRT of training first-level facility health workers in a clinical algorithm for the management of malaria, and in microscopy for malaria diagnosis. A 34 percent reduction in antimalarial prescriptions was seen in health centres with the algorithm and microscopy training, in comparison to those where staff were trained with the algorithm only. The reporting of this CRT resulted in an unclear risk of bias in terms of the extent of adjustment for baseline differences between areas, although other categories were at low risk of bias.

Effect on quality

The evidence from two studies that reported the effect of a health service delivery strengthening intervention on the quality of malaria treatment among febrile children aged <5 years was mixed:

- Gilroy and colleagues (2004) reported on a CRT where head nurses from community health centres received training in the WHO's IMCI guidelines. They reported no significant improvement in antimalarial drug counselling (dosage, timing and duration of treatment course). The study was subject to high risk of bias in terms of important differences in socio-economic status between groups at baseline that were not adjusted for in analysis (Gilroy et al., 2004).

- Zurovac and colleagues (2011) reported on a CRT of text message reminders for the management of non-severe malaria to first level facility health workers. The authors reported a >20 percent improvement in the appropriate management of non-severe malaria, as defined by the proportion of consultations with febrile children under five years in which tasks were completed. The results were sustained up to at least six months after the intervention. This study was at low risk of bias.

The evidence from a CRT training lady health visitors in the WHO's IMCI nutrition counselling module showed mixed results for the quality of nutritional counselling, with significant improvement in some but not all measures. Action tasks of weighing, plotting growth and checking feeding practices were between 1.5 and 10 times more likely in the intervention group. This study was at low risk of bias (Zaman et al., 2008).

A multi-country CRT study investigated the effectiveness of access to the WHO's Reproductive Health Library electronic resources on obstetric practices. The authors reported no significant difference in the mean change in the proportion of women experiencing any of the six practices. This study was at low risk of bias (Gulmezoglu et al., 2007).

Effect on clinical outcomes and/or mortality

Zaman and colleagues (2008) also measured anthropometric outcomes in their CRT of training lady health visitors, which was at low risk of bias. They reported improved weight- and height-for-age in 6-18 month olds (although neither intervention nor control group had <−2 z-scores), and no significant improvement in height-for-age (again, neither group was stunted).

Table 3.4: Outcomes of studies with interventions classified as addressing improving the service provided by front-line workers during their interaction with users only (Category A)

Reference and study period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post-intervention	Outcome category			Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Quality	Coverage	Clinical or Impact			
Gilroy et al. (2004) 2001	Mali; rural; District (n=1)	IMCI: case management improvement element	CRT (n=10 community health centres)	6 weeks	✓			Antimalarial drug counselling (dosage, timing and duration of treatment course) Mean of a 10-point composite score	Observations of case management of febrile <5 years prescribed antimalarials (n=364)	Intervention centre mean score 1.47 higher (95%CI -0.25, 3.19) p>0.05
Gulmezoglu et al. (2007) 2001	Peri-urban Thailand and urban Mexico; District (n=1 Mexico, n=1 Thailand)	WHO RH Library: Maternity hospital clinical staff access to WHO's RH Library electronic resources	CRT (n=18 hospitals Thailand; 22 hospitals Mexico)	4-6 months	✓			6 obstetric practices. Mean proportion of women experiencing any of 6 practices adjusted for the proportion at baseline. Results country-specific	Hospital records (n varied by obstetric practice: min. n=20, max. n=13,037)	Difference between baseline and endline mean proportion of women experiencing practices not significantly different in Mexico or Thailand

Reference and study period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post-intervention	Outcome category			Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Quality	Coverage	Clinical or Impact			
Ngasala et al. (2008) 2003-4	Tanzania; rural; District (n=2)	Malaria treatment and diagnosis: HW trained in malaria management only or additionally in microscopy diagnosis versus no special training	CRT (n=16 health facilities)	?		✓		% Antimalarial prescriptions among febrile children <5 years	Health facility records of consultations with febrile <5 years (n=3131)	61% (95%CI 51.6, 70.3%) algorithm + microscopy vs 95.2% (95%CI 85.7, 100%) algorithm only vs 99.4% (95%CI 90.9, 100%) no special training
Zaman et al. (2008) ?	Pakistan; urban; District (n=1)	IMCI nutrition: IMCI nutrition counselling module to train lady health visitors	CRT (n=40 health facilities)	Quality: ? Clinical: 14, 45 and 145 days	✓		✓	Quality: Nutrition communication and case management (10 communications, 4 feeding questions and 3 assessment tasks).	Quality: Observations of consultations with children aged 6-18 months of age (n=105) Clinical: Home	Quality: Significant improvement in 5/10 communication, 2/4 feeding questions and 3/3 action tasks. Action tasks: weigh child

Reference and study period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post-intervention	Outcome category			Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Quality	Coverage	Clinical or Impact			
								Odd ratios of tasks Clinical: WAZ, HAZ, WHZ; means of WAZ, HAZ, WHZ	visits (n=170 day 14, n=168 day 45, n=142 day 145)	(OR=1.52; 95% CI 0.50, 4.64); plot growth (OR=7.05; 95%CI 1.16, 42.7); check feeding vs recommendations (OR=10.4; 95%CI 1.91, 56.8) Clinical: WAZ: -1.20 vs -1.72 (p=0.012); HAZ: -1.60 vs -1.70 (p=0.6); WHZ: -0.30 vs -0.80 (p=0.005)

Reference and study period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post-intervention	Outcome category			Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Quality	Coverage	Clinical or Impact			
Zurovac et al. (2011) 2009	Kenya; rural; District (n=11)	SMS malaria: Health workers sent text message reminders for the management of non-severe malaria	CRT (n=11 districts)	Immediately and 6 months	✓			Management non-severe malaria with AL (completion of 4/4 case management and 4/6 dispensing and counselling tasks) Difference of differences in mean proportion of correctly managed cases	Interviews with carers of febrile children <5 year on health worker case management, dispensing and counselling behaviours (n=861 baseline, pre-intervention n=749 immediately post-intervention and n=659 six months post-intervention)	Difference of differences 23.7% (95%CI 7.6, 40.0%) immediately, and 24.5% (95%CI 8.1, 41.0%) 6 months later (p<0.05)

(B) Improving the service provided by front-line workers during their interaction with users, with additional strengthening at the service delivery level

Details of studies in this category are shown in Table 3.5.

Effect on coverage

Ayieko and colleagues (2011) in their CRT of supporting the implementation of paediatric guidelines, reported improvements of over 19 percent in the coverage of provider-initiated HIV testing and counselling for patients where HIV status was unknown. Recording immunisation status also improved by over 25 percent, but a 28 percent improvement in Vitamin A on hospital admission was not statistically significant. This study was at low risk of bias.

Effect on quality

The evidence on quality available from the four studies in this intervention category suggests a positive association between health service delivery strengthening and quality of care:

- Jennings and colleagues (2010) reported on a CRT of job aids for antenatal counselling. The authors reported that nurses and midwives in intervention maternity units delivered almost 20 percent more messages on average than in control facilities. However, this study was at high risk of bias in terms of differences in the characteristics of the pregnant women between areas at baseline that were not adjusted for in the analysis.

- Rowe and colleagues (2009) reported on a CRT of enhanced supervision after IMCI training. The study was implemented over four years rather than the intended one year, and as such also presented per-protocol (PP) analyses, as contamination of comparison areas was reported to be likely. Intention-to-treat results showed no significant effect. However, in PP analysis, the difference of differences in the adherence index showed a 15.5 percent increase for observed consultations carried out by health workers who had received supervision supported by the study, in comparison to those who had received usual supports. This study had unclear risk of bias for some aspects.

- Bradley and Igras (2005) reported on a non-randomised cluster study of a cycle of quality improvement self-assessments carried out in health centres. This study reported positive associations with many aspects of the quality of the management of childhood illness, although taking in ill children did not show as much improvement as other tasks. Because of its non-randomised design, this study was at high risk of bias on most categories of the EPOC checklist.

- Ayieko and colleagues (2011) also reported on the quality of case management of childhood illness, with a 29 percent improvement in mean scores in intervention hospitals.

Table 3.5: Outcomes of studies with interventions classified as improving the service provided by front-line workers during their interaction with users with additional strengthening at the service delivery level (Category B)

Reference and study period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category			Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Quality	Coverage	Clinical or Impact			
Ayieko et al. (2011) 2006-08	Kenya; rural; District (n=4)	Implementation of guidelines: *Health service delivery level (II) Training* Staff were trained in ETAT+ guidelines during 17 lectures and practicals over 5.5 days and provided with job aids *Health service delivery level (II) Management and Supervision* External enhanced supportive supervision was planned every 2-3 months, and a full-time in-facility facilitator was nominated. 6-monthly health facility surveys with face-to-face feedback, over a period 18 months	CRT (8 district hospitals)	18 months	✓	✓		Quality: Mean child-level score of % counselling tasks performed Difference in differences in mean score. Coverage: Provider-initiated HIV testing and counselling Documented immunisation status Vitamin A administered on hospital admission Difference in differences in prevalence	Hospital records for paediatric admissions <5years (n=2,135 baseline, n=2,305 endline)	Quality: Mean score 0.29 (95%CI 0.05, 0.54) higher in intervention hospitals Coverage: HIV testing 19.4% (95%CI 12.3, 26.4%); immunisation status 25.8% (95%CI 7.3, 44.4%); Vitamin A 28.3% (95%CI -7.1, 63.6%)

Reference and study period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category			Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Quality	Coverage	Clinical or Impact			
Bradley and Igras (2005) 1999-2001	Guinea and Kenya; mixed urban and rural; District (n=8 Guinea, n=4 Kenya)	Health centre quality improvement: *Health service delivery level (II) Training* Staff underwent training on supervision, communication, counselling, immunisation, and infection prevention *Health service delivery level (II) Management* Staff participated in sequential self-administered quality improvement assessment sessions over 2-3 days, followed by planning sessions for an action plan. Process repeated every 3-4 months over 15 months	Non-randomised cluster design (4 districts Kenya; 8 districts Guinea)	15 months	✓			Quality of management of childhood illness based on 45-item checklist (reflects IMCI protocols) Percentage of consultations in which each behaviour observed	Observed consultations with sick and well children (n=160 Guinea, n=160 Kenya)	All items for well children improved. Privacy, interpersonal skills, information improved. Most physical examination and prescribing tasks improved. No improvement in history taking

Reference and study period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category			Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Quality	Coverage	Clinical or Impact			
Jennings et al. (2010) 2008	Benin; rural; Region (n=2)	ANC job aids: *Health service delivery level (II) Training* Training for 3 days in the use of job aids to guide counselling during antenatal visits *Health service delivery level (II) Management* Maternity units were also encouraged to form teams to identify organisational changes that were needed	CRT (14 maternities)	3 weeks	✓			Mean percent of messages provided to pregnant women by nurses and midwives Difference in differences of the mean percentage of messages provided	Observed consultations (n=330 baseline, n=356 endline)	Intervention maternity units 19.6% (95% CI 12.2, 26.9%) increase in mean % of messages provided
Rowe et al. (2009) 1999-2004	Benin; mixed urban and rural; District (n=8)	Post IMCI supervision: *Health service delivery level (II) Training* 11-day IMCI case management improvement course *Health service delivery*	CRT (8 districts)	1-37 months	✓			Child-level adherence index of % of needed IMCI tasks performed Difference of differences in needed tasks	Observed consultations with children <5 years (n=228 baseline 1999; n=225 survey 1 2001; n=77 survey 2	15.5% (95%CI 9.0, 22.0%) increase in adherence index [per protocol]

Reference and study period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category			Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Quality	Coverage	Clinical or Impact			
		level (II) Supervision Enhanced supportive supervision of 2 supervision visits every 3 months planned. Job aids. Supervision of supervisors was also included, and face-to-face feedback by supervisors to facilities						performed	2002; n=156 survey 3 2004). Per protocol, n= 430 pre-intervention (1999-2001); n=814 post-intervention (2002-04)	

(C) Improving the service provided by front-line workers during their interaction with users, in addition to strengthening at the service delivery, community or household, or health sector policy levels

Details of studies in this category are shown in Table 3.6.

Effect on coverage

Winch and colleagues (2003) reported on a CRT of malaria case management training for village drug kit managers:

- The authors reported over 50 percent more children diagnosed with malaria who received CQ at home according to IMCI guidelines in intervention areas.
- However, completion of referral among children referred by village drug kit managers to health centres, measured by carer recall, showed no difference between areas.
- This study was of unclear risk of bias on some items, and of high risk of bias in terms of differences in caregiver characteristics at baseline that were not adjusted for in the analysis.

Armstrong Schellenberg and colleagues (2004a, b) reported on a quasi-experimental study to evaluate IMCI:

- The authors reported a positive but non-statistically significant 5 percent increase in coverage of care seeking from an appropriate provider, and a positive but non-statistically significant 9 percent increase in the prevalence of exclusive breastfeeding.
- However, appropriate treatment of children diagnosed with pneumonia or anaemia was 35-40 percent higher, and that of children diagnosed with malaria was over 60 percent higher during consultations with health workers in districts where IMCI implementation had taken place.
- Due to its non-randomised design, this study was at high risk of bias across many of the EPOC checklist items.

Arifeen and colleagues (2004a, b) reported on a CRT to evaluate IMCI:

- The authors reported a significant 10 percent improvement in care seeking and in exclusive breast feeding in intervention areas.
- Treatment according to guidelines for children diagnosed with a priority illness (pneumonia, diarrhoea with dehydration, dysentery, malaria or measles) was almost 60 percent higher in intervention areas.
- This study was at low risk of bias.

Effect on equity

One study with high risk of bias measured equity (Tanzanian IMCI). The results showed a positive association between strengthening health service delivery and equity of treatment of underweight, stunting and fever, as well as measles vaccination and bednet use, using a concentration index of household-level asset ownership, household characteristics, education and income. Although improvements in equity were seen, malnutrition outcomes remained pro-rich, with more children in the poorest quintile versus the least poor quintile classified as

malnourished (Masanja et al., 2005, a satellite study of Schellenberg et al., 2004a, b).

Effect on quality
The evidence from all three studies in this category showed a positive association between strengthening health service delivery and the quality of case management in children:

- Winch and colleagues (2003) reported that the quality of antimalarial treatment counselling, as defined by completion of three tasks, was over seven times more likely in children from intervention areas with access to trained village drug kit managers.

- Armstrong Schellenberg and colleagues (2004a) reported improvements in 25 of 29 indicators of assessment, classification, treatment, counselling and communication in intervention areas implementing IMCI.

- Arifeen and colleagues (2004a, b) reported that correct case management based on the mean of child-specific indices was 53 percent improved in intervention areas implementing IMCI.

Effect on nutritional outcomes
Both studies in this category that evaluated nutritional outcomes in young children were IMCI evaluations. In Tanzania, a 30 percent improvement in low weight-for-age was reported. No improvement in wasting was reported by either study. No improvement in stunting was reported in Tanzania, and although the intervention area children had 7 percent lower prevalence of stunting in Bangladesh, this remained extremely high at over 50 percent.

Effect on mortality
The two IMCI evaluations in Bangladesh and Tanzania were the only two studies of the 12 included in this review that reported outcomes of mortality. Neither reported a statistically significant impact on survival of children aged under 5 years. However, non-statistical declines in point estimates of mortality were seen: 27.0 per 1,000 children in intervention districts and 28.2 in comparison districts in Tanzania, and 27.0 per 1,000 children in IMCI intervention areas and 31.2 in comparison areas in Bangladesh.

Table 3.6: Outcomes of studies with interventions classified as improving the service provided by front-line workers during their interaction with users in addition to strengthening at the service delivery, and/or community/household or health sector policy levels (Category C)

Reference and Study Period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category				Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Equity	Quality	Coverage	Clinical or Impact			
(Arifeen et al., 2004) 2001-07	Bangladesh; rural; District (n=1)	IMCI: *Health service delivery level (II) Training* Training 11-day IMCI case management improvement course, job aids and supplies *Health service delivery level (II) Supplies* Provision of essential drugs and supplies and introduction of facility supply tracking and reporting system *Health service*	CRT (n=20 catchment areas)	Quality: 18 months Coverage: 18 months Impact: 5 years		✓	✓	✓	Quality: quality of management of childhood illness based on index of correct case management (0-100) Coverage: % seeking care for child illness from appropriate provider (2003) % aged <6 months exclusively breastfed	Quality: Observation of consultations with children <5yrs (n=342) Coverage: Household survey 2003 (n=944) Observation of consultations with children <5yrs diagnosed with a priority illness (n=108) Impact: Household census 2006-7 (n=68,945?) anthropometric sub sample	Quality: mean index intervention 73, comp 17 (p<0.001) Coverage: Appropriate provider 18% intervention, 8% comp. (p=0.05) Exclusively breastfed 75.5% intervention 65.3% comp (p=0.011) Appropriate treatment for priority illness (79% vs 21% p<0.001)

Reference and Study Period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category				Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Equity	Quality	Coverage	Clinical or Impact			
		delivery level (II) Management and Referral Referral forms were linked to an introduced referral system, with sensitisation activities for referral facility staff. *Health service delivery level (II) Shortage and Distribution of Staff* Guidelines for the management of pneumonia at first-level facilities was amended to allow treatment prior to referral							(2006-7) % diagnosed with priority illness receiving appropriate treatment Clinical and impact: HAZ 2-5 years, WHZ <2 years (2006-7) <5 years mortality rate 7 days - <5 years mortality per 1,000 children and rate ratio (2006-7)	(n=4,400)	Clinical and impact: WHZ12.6% intervention 14.3% comp (p=0.712); HAZ 50.4% intervention 57.1% comp (p=0.029) 7days - <5 year mortality rate 27.0 intervention 31.2 comp. (p=0.31)

Reference and Study Period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category				Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Equity	Quality	Coverage	Clinical or Impact			
		of severe cases, and later a new cadre of community health worker introduced to manage uncomplicated pneumonia at the community level. *Community/Household level (I)* Community mobilisation included training of local imams and the use of community theatre to encourage health seeking									

Reference and Study Period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category				Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Equity	Quality	Coverage	Clinical or Impact			
Armstrong Schellenberg et al. (2004a, b); Masanja (2005) 2000-02	Tanzania; rural; District (n=4)	IMCI: *Health service delivery level (II) Training* 2 of the 3 IMCI elements implemented, with 11 day IMCI case management improvement course *Health sector policy and strategic management level (III)* Increased autonomy for district council health management teams on budgeting and	Quasi-experimental cluster design (n=4 districts)	Quality: 3 years Equity, coverage and impact: 5 years	✓	✓	✓	✓	Equity: 10 outcomes (malnutrition, malaria treatment, immunisation, insecticide treated bed net, care seeking). Concentration index across quintiles of household wealth. Quality: quality of management of childhood illness based on 29 indicators of assessment, classification, treatment, counselling and	Equity: Household surveys in 1999 (n=2006) and 2002 (n=1924), and difference of differences in concentration index between periods for each of 10 outcomes. Quality: Observations of consultations with children <5yrs (n=419) Coverage and clinical: *H*ousehold surveys baseline 1999, endline 2002 (n=2,006 children in 1,321 households; n=1,932 children in 1,341	Equity: Improvement in equity seen in all outcomes over time, significantly higher increase in IMCI areas for underweight and stunting, measles vaccination, ITN use, and malaria treatment Malnutrition outcomes remained pro-rich Quality: 25/29 indicators significantly better in intervention

Reference and Study Period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category				Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Equity	Quality	Coverage	Clinical or Impact			
		planning, linked to availability of district level burden of disease data, a budget modelling tool, and simulated basket funding – all linked to TEHIP.							communicatio n. Coverage: % seeking care for child illness from appropriate provider % aged <4 months exclusively breastfed % diagnosed with pneumonia, malaria, or anaemia receiving appropriate treatment	households) Observations of consultations with children <5yrs diagnosed with pneumonia (n=111), malaria (n=304), or anaemia (n=84) *Impact:* Demographic surveillance of district-specific mortality rates between mid-2000 and mid-2002	Coverage: Care from appropriate provider 38.2% intervention 32.3% comparison p=0.45 Exclusively breastfed 22.7% intervention 32.1% comp. p=0.69 Appropriate treatment for pneumonia (75% v 40% p<0.01); malaria (88% v 25% p<0.001), or anaemia (44% v 4% p=0.01) *Clinical and impact:* WAZ 23.3% intervention

Reference and Study Period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category				Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Equity	Quality	Coverage	Clinical or Impact			
									Clinical and Impact: WAZ, HAZ, WHZ <5 years mortality per 1,000 children and rate ratio		53.4% comp. p=0.03, HAZ 50.7% intervention 40.0% comp. p=0.07, WHZ 6.8% intervention 5.6% comp. p=0.43 <5yr mortality 27.0 intervention 28.2 per 1000 child years. Rate ratio 0.87 p=0.28, no cluster p=0.004
Winch et al. (2003) 2001	Mali; mixed urban and rural; District (n=1)	Community treatment of malaria: *Health service delivery level (II) Training* Training in age-specific dosing	CRT (n=10 health zones)	2-10weeks		✓	✓		Quality: Quality of antimalarial counselling, i.e. telling the caregiver the dose to administer the	Caregiver recall (n=286)	Quality of all 3 antimalarial counselling tasks: (OR=7.7; 94%CI 1.4, 40.4) Coverage: 72% intervention vs

Reference and Study Period	Country, urban/ rural and scale	Intervention	Study design (n= unit of random or non-random intervention allocation)	Timing of evaluation post intervention	Outcome category				Outcome description and outcome measure, or measure of effect	Source of outcome information and sample size	Effect (95% CI) (p value)
					Equity	Quality	Coverage	Clinical or Impact			
		of CQ over 5 days. *Health service delivery level (II) Management and Referral* Raised awareness with head nurses of intervention area community health centres (see study above by Gilroy et al., 2004) *Community/ household level (I)* Included village meetings to encourage use of the village drug kits							child, the number of days of treatment, and showing the caregiver how to administer the drug. Odds ratio % of consultations where all 3 drug counselling tasks were completed. Coverage: % diagnosed malaria cases receiving CQ at home according to guidelines Odds ratio % referrals completed		22% comparison; p<0.001 completion of referral (OR=0.14; 95%CI 0.01, 1.93)

4. Findings from other systematic reviews of interventions to improve health service delivery in LMICs

4.1 Overview

We identified 12 systematic reviews (one of which is still in process) related to interventions implemented to target constraints at the health service delivery level. These reviews were identified as part of the literature review for the background section of this report, and updates that authors were aware of through their networks. They are not the result of a comprehensive search and this summary is therefore not a comprehensive summary of systematic reviews in this area. Some characteristics of these systematic reviews are summarised in Table 4.1 and further details shown in Appendix 9.

To summarise some of the key messages from these systematic reviews, we have grouped them according to the constraint within Hanson and colleagues' (2003) health service delivery level that the interventions aimed to address:

- Four reviews addressed the shortage and distribution of appropriately qualified staff;
- Three addressed weak technical guidance;
- One addressed weak programme management;
- One addressed weak supervision;
- Three reviews addressed weak technical guidance, as well as weak programme management and weak supervision.

4.1.1 Shortage and distribution of appropriately qualified staff

The four systematic reviews addressing this constraint included evidence from LMICs, although the majority of evidence was drawn from high income countries (23/89 studies from LMICs). All restricted study design inclusion to robust epidemiological designs (RCT, non-randomised CT, CBA, ITS). Two addressed interventions aiming to retain health workers in LMICs or recruit and retain health workers in rural areas of LMICs (Grobler et al., 2009, Penaloza et al., 2011). Two focused on the training of non-professional health workers (traditional birth attendants – TBAs - and lay health workers) for maternal, neonatal and child health (Lewin et al., 2010, Sibley et al., 2007).

- There was little to no robust evidence on recruiting or retaining health workers in LMICs, with only one study meeting inclusion criteria over the two reviews addressing these questions;
- The evidence base for lay health workers, including TBAs, was wider;
- Some positive associations were reported of using lay health workers with outcomes, including 25 percent-30 percent reductions in neonatal and child mortality from seven studies across two reviews, although the quality of the evidence was graded low.

4.1.2 Weak technical guidance

The three systematic reviews addressing this constraint included evidence from LMICs (52/121 studies). One review restricted study designs to robust designs (RCTs, CBAs and ITS studies) (Opiyo and English, 2010), one to randomised controlled trials (Forsetlund et al., 2009), while the third had no restrictions on design as long as outcome data was included (van Lonkhuijzen et al., 2010). One investigated the effect of educational meetings and workshops on professional health workers behaviour (Forsetlund et al., 2009), and the others investigated the effect of training programmes aimed at improving care of severely ill newborns and children, and improving emergency obstetric care (Opiyo and English, 2010, van Lonkhuijzen et al., 2010).

- There were few robust studies from LMICs;
- There was some evidence to suggest small improvements in health worker knowledge related to training programmes aimed at improving emergency obstetric care among skilled birth attendants, but that this for the most part did not translate to improved behaviour;
- There was some evidence to suggest small improvements in health worker behaviour for newborn resuscitation following in-service training;
- Evidence suggests that educational meetings may have small effects on improved behaviour, but that these are unlikely to be sufficient to change complex behaviour;
- There was low-quality evidence from one study to suggest that in-service training on the treatment of seriously ill newborns did not improve neonatal mortality.

4.1.3 Weak programme management

One review assessed the effects of audit and feedback on the practice of healthcare professionals, restricted to RCTs (Jamtvedt et al., 2003):

- The majority of evidence was from high-income countries (4/118 from LMICs);
- Small effects on knowledge and behaviour were found, and these were largest when the baseline levels were low;
- There is no evidence to suggest the use of audit and feedback as a mandatory intervention to improve health worker behaviour.

4.1.4 Weak supervision

One review assessed the evidence of the effect of managerial supervision of health workers on the quality of primary health care in LMICs, restricted to robust designs (RCT, CBA, ITS) (Bosch-Capblanch et al., 2011):

- Nine studies from LMICs in Africa, Asia and Latin America were included;
- This review suggested that there were some small benefits to health worker knowledge and practice, but not enough high quality evidence to advocate any particular form of supervision;
- There was evidence to suggest that monthly versus quarterly supervision did not show improved results.

4.1.5 Weak technical guidance together with weak programme management and weak supervision

Three systematic reviews addressed at least two of the constraints of weak technical guidance, weak programme management and weak supervision (Oliveira-Cruz et al., 2003, Rowe et al., 2012, Smith et al., 2009). Two were general in their approach (Oliveira-Cruz et al., 2003, Rowe et al., 2012), and one focused on malaria (Smith et al., 2009). All were restricted to studies from LMICs. Study design inclusion criteria varied from any design to more robust methodologies. One review is ongoing and results presented are preliminary (Rowe et al., 2012).

- There were few robust studies and some reviews included a wide range of study designs;
- In-service training interventions to improve prompt and effective treatment of fever in children under five years of age on their own did not show positive or large effects, although combining training with job aids and frequent supervision showed better results;
- The authors highlighted that evidence on the sustainability of any changes in knowledge or practice was lacking in relation to in-service training to improve malaria case management;
- Interventions to improve health worker skills (training, supervision, quality assurance) showed mixed effects on improved health worker knowledge;
- Training interventions targeting prescribing practices were generally positive;
- Quality assurance interventions improved waiting times on the whole;
- Management strengthening interventions did not tend to report outcomes of service utilisation or coverage; however, it was suggested that process indicators such as rational use of resources, data collection and use, and staff morale had improved;
- The authors highlighted that the evidence was insufficient with respect to equity and whether interventions were successfully reaching the poor and disadvantaged;
- The authors summarised the effects of the absolute changes in process outcomes of health worker practice using a median effect size, and many (40 percent of studies) showed a small effect size of <10 percentage point increases;
- Interim results suggest that interventions with numerous components (e.g. training + job aids + supervision + community activities) have improved effects and larger median effect sizes in comparison to studies with only one component;
- Studies that also include a community activity component also seem to have improved effects.

4.2 Comparison of findings

We can make a number of observations based on the evidence synthesised in this review, and on that available from other systematic reviews on linked topics.

- **Evidence base for efficacy and effectiveness:** This review, and others, have emphasised the relative scarcity of studies carried out using robust study designs to generate evidence of efficacy or effectiveness in LMICs. Furthermore other reviews which undertook assessment of the quality of the evidence for impact (e.g. lay health workers) concluded that the quality of the evidence for impact is generally low to very low, as defined by the GRADE criteria. Other reviews which had more inclusive inclusion criteria for study design, commented on the difficult of comparing studies due to poor study quality and variability in methods of measuring outcomes.

- **Outcomes:** There is a focus on outcomes of health worker knowledge, with fewer assessing health worker behaviour and very few assessing outcomes of survival impact.

- **Technical guidance alone is insufficient:** Although the authors of most other reviews identified did not explicitly classify interventions by their components, other reviews did emphasise that simple training interventions often did not have positive or large effects on behaviour, and that interventions such as educational meetings alone were insufficient to change complex behaviour. The authors of other reviews highlighted that interventions including training were more effective when training was included as part of a package of interventions, combined with, for example, enhanced supportive supervision.

- **Shared characteristics of success:** Our review also highlighted three other characteristics in addition to moving beyond technical guidance alone. These included: 1) sustained interaction; 2) local ownership through a focal point person; and 3) feedback cycles. These findings were not specifically drawn out in the results or conclusions of other systematic reviews.

Table 4.1: Overview of systematic reviews of interventions to improve health service delivery in LMICs

Reference	Title	Setting (Countries)	Study designs	Number of studies	Participants	Interventions	Outcomes
Bosch-Capblanch (2011)	Managerial supervision to improve primary healthcare in low- and middle-income countries	LMIC (Africa, Asia, Latin America)	RCT, CBA, ITS	9	Health workers in primary healthcare in LMICs	Supervision: site visits from a central level of health system, plus at least one supervisory activity	Adherence to guidance
Forsetlund et al. (2009)*	Continuing education meetings and workshops: effects on professional practice and health care outcomes	HICs and LMICs. 12 studies in Africa, Asia, Latin America	RCT	81	Health professionals in high or LMICs	Educational meetings (conferences, lectures, workshops, seminars, symposia and courses)	Compliance with desired practice
Grobler et al. (2009) *	Interventions for increasing the proportion of health professionals practising in underserved communities	HICs and LMICs	RCT, CBA, ITS	None of the 90 studies met study design criteria	All qualified healthcare professionals	Interventions to increase proportion of healthcare professionals working in rural areas	Recruitment; Retention
Jamvedt et al. (2003) *	Audit and feedback: effects on professional practice and heath care outcomes	HICs and LMICs. 4 studies were in Africa or Asia	RCT	118	Healthcare professionals	Audit and feedback	Compliance with desired practice

Reference	Title	Setting (Countries)	Study designs	Number of studies	Participants	Interventions	Outcomes
Lewin et al. (2010) *	Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases	HICs and LMICs.16 trials from 10 countries in Africa, Asia, Latin America and the Caribbean	RCT	82	Lay health workers (without formal professional certification)	Training of lay health workers in maternal and child health and infectious diseases	Immunisation uptake; Breastfeeding; Health seeking; TB; Under-five mortality
Oliveira-Cruz et al. (2003)	Approaches to overcoming constraints to effective health service delivery	LMICs in Africa, Asia, Latin America and the Caribbean	Any	56	Health workers within the close-to client health system	Management system strengthening, training and supervision, introduction of quality assurance systems	Efficient use of inputs; provider knowledge; client perception of service quality; utilisation; coverage; morbidity or mortality
Opiyo and English (2010) *	In-service training for health professionals to improve care of the seriously ill newborn or child in low and middle-income countries	LMIC. 1 study each from Kenya and Sri Lanka	RCT, CBA, ITS	2	Health professionals responsible for management and care of seriously ill newborn/child	In-service training for the care of seriously ill newborns or children	Adequate initial resuscitation; Assessment of breathing and newborn care practices

Reference	Title	Setting (Countries)	Study designs	Number of studies	Participants	Interventions	Outcomes
Penaloza et al. (2011) *	Interventions to reduce emigration of health care professionals from low- and middle-income countries	LMIC. 1 study from Philippines	RCT, non-randomised CT, CBA, ITS	1	Health professionals, nationals of a LMIC whose graduate training was in an LMIC	Intervention in the source, recipient or both countries that could impact on number of professionals emigrating	Number of nurses migrating from Philippines to the USA annually
Rowe et al. (2012) †	How can we achieve and maintain high-quality performance of health workers in low-resource settings?	LMIC	Studies with comparison groups	>500	Health care workers including community and lay health workers	Interventions to improve health worker behaviour	Health worker practices; patient outcomes; costs
Sibley et al. (2007) *	Traditional birth attendant training for improving health behaviours and pregnancy outcomes	LMIC. All studies from rural LICS in Asia, Africa and Latin America	RCT, CBA, ITS	4	Traditional birth attendants	Traditional birth attendant training	Maternal morbidity; Maternal; Perinatal; Neonatal mortality
Smith et al. (2009)	Provider-side interventions to improve malaria treatment	LMIC. 13 African malaria-endemic countries	RCT, ITS, CBA uncontrolled BA, controlled after	16	Providers responsible for dispensing antimalarials in malaria endemic LMICs	Interventions targeting provider knowledge or behaviour	Prompt and effective treatment for malaria

Effectiveness of interventions to strengthen national health service delivery on coverage, access, quality and equity in the use of health services in low and lower middle income countries

Reference	Title	Setting (Countries)	Study designs	Number of studies	Participants	Interventions	Outcomes
van Lonkhuijzen et al. (2010)	Interventions to improve quality of emergency obstetric care	LMICs (Asia, Africa, Latin America)	Any	38	Skilled birth attendants (WHO definition)	Postgraduate training programmes aimed at improving emergency obstetric care	Knowledge and skills

* Summaries for these reviews are available from the SUPPORT Collaboration: www.supportsummaries.org;

† This systematic review is ongoing and reported findings are interim results.

5. Discussion

5.1 Summary of findings

Findings from this review suggest that studies of interventions which strengthened other elements of the health service delivery in addition to technical guidance, as well as community mobilisation and interventions at the health sector policy and strategic management level, showed more consistent improvement on quality of care and counselling, than those using technical guidance alone.

Supply-side interventions that appeared to have no, little, or substantially mixed effects on quality of care included:

- providing access to guideline information electronically;
- providing access to and training in algorithms for malaria management only, without the addition of training in microscopy;
- providing training for antimalarial drug counselling in the absence of other supportive interventions;
- providing training for nutritional counselling in the absence of other supportive interventions.

Supply-side interventions that appeared to have a positive effect on quality of care included:

- text message reminders (with motivational quotes) for malaria case management;
- training for malaria case management, when combined with community awareness, supervision and referral mechanisms;
- job aids for antenatal counselling when combined with supervision and a focus on institutional adaptations required to incorporate the use of job aids;
- IMCI training, when implemented in combination with enhanced supervision that incorporated training of supervisors, job aids, use of data and face-to-face supportive supervision, in repeated cycles of assessment, examination/feedback and planning (Plan-Do-Study-Act-Plan cycle);
- implementation of guidelines, when delivered using training, enhanced supportive supervision, a focal person to troubleshoot problems on site, and repeated progress surveys with face-to-face feedback and planning sessions (Plan-Do-Study-Act-Plan cycle);
- quality improvement, when combined with training, supervision, repeated progress surveys with time frames and named individuals identified against decisions/plans made during face-to-face meetings with all health facility staff, and district level representation (Plan-Do-Study-Act-Plan cycle);
- implementation of full IMCI guidelines, incorporating training, supervision and discussion of how to overcome barriers to implementation, wider health system strengthening at the health sector policy and strategic management level, and community mobilisation.

5.2 Characterising interventions

We used a framework suggested by Hanson and colleagues (2003) to characterise included interventions to improve health service delivery by front-line workers in LMICs. We suggest that this type of approach is useful, not only for synthesising findings across studies (Anderson et al., 2011, Shepperd et al., 2009), but also to help inform policy makers in deciding what elements of an intervention are applicable and sufficient in their context.

Characterising the interventions in this way illuminates the 'black box' of strengthening health service delivery (Murray and Frenk, 2000). We suggest that, along with detailed reporting of context and implementation processes, as recommended by the MRC in its guidelines for the evaluation of complex interventions (Craig et al., 2008), such characterisation is an essential first step for moving from questions of effectiveness to health system and policy research questions of 'what happened, how, and why' or why not (Sheikh et al., 2011).

Our experience in conducting this review suggests that future evidence synthesis of large-scale delivery of complex public health interventions would benefit from improved consistency of reporting of the implementation process by included studies (Craig et al., 2008, Oakley et al., 2006); recommendations for reporting are available from the Standards for Quality Improvement Reporting Excellence (SQUIRE) guidelines, which suggest a 19-point checklist (Davidoff et al., 2009).

Some recommendations to highlight include:

- description of the local problem: the nature and severity of the specific local problem or system dysfunction;
- intended improvement: the specific aim of the proposed intervention; and who and what triggered the decision to make changes, and why now;
- specification of the way the elements of the local care environment considered most likely to influence change were identified and characterised;
- intervention: description of the component parts in sufficient detail that others can reproduce it;
- description of the planning and implementation of the intervention;
- evaluation methods: description of plans for assessing how to measure dose or intensity of exposure and how well intervention was implemented;
- logic model: description of the mechanisms by which the intervention components were expected to cause changes, and plans for testing whether those mechanisms were effective.

5.3 Potential pathways of effect and mapping the evidence base

A simple framework of the potential pathway between health system strengthening and outcomes was presented in Figure 2.2. The findings from the 12 included studies were used map out what sections of the framework have evidence available from this review (Figure 5.1). This framework is still simplistic and addresses only health service delivery (Hanson et al., 2003); certain pathways are not addressed by this review and are shown in the figure as dashed grey lines.

- The evidence base for quality outcomes is the largest, with the majority of studies reporting this outcome (green line).
- This review found no or sparse evidence of interventions delivered at scale and evaluated robustly for outcomes of access or equity (red lines).
- Findings from the review showed some mixed evidence on a wide range of coverage indicators, which vary in how closely they relate to measures of disease prevalence or incidence (amber lines).

Evidence for survival impact was limited to two studies, both of Integrated Management of Childhood Illness, which showed positive but not statistically significant associations with mortality in under-fives.

Logic models have been used in the assessment of other health strengthening interventions (Adam et al., 2012). Interesting examples include work by Huicho and colleagues (2010) in relation to the recruitment and retention of health workers in underserved (rural/remote/poor) areas. van Olmen and colleagues (2012a) have drawn specifically on theory-driven evaluation to guide the development and use of logic models for health systems research. They have used this approach in evaluation of interventions - for example work undertaken by van Belle and colleagues (2010) on evaluation of an intervention to improve the continuity of care in reproductive health in three West African settings.

We also suggest that it may be of use to policy makers to map out assumptions about inputs and processes (including situation analysis, implementation and others recommended in the SQUIRE guidelines and highlighted above), that can be monitored during implementation, as well as guiding the evaluation of a programme to strengthen health service delivery (World Health Organization, 2010a).

Gaps highlighted in this review from the mapping exercise to the logic model, are likely to require a number of approaches, ranging from robust effectiveness evidence on survival impact to contextual detail and a range of methodological approaches. Some recommendations have been recently outlined by Gilson and colleagues, including a focus on strategies as approaches to research that use the most appropriate method for the research question (Gilson, 2012). The handbook points to seven suggested strategies and includes examples of each of these.

An overview of the seven suggested strategies is shown below:

1. Cross-sectional perspectives: focused on one point in time and may use qualitative, quantitative or mixed methods, including triangulation from multiple data sources. These require relatively few resources.
2. The case-study approach: exploring within the real-world setting and using multiple sources of evidence. This approach is relevant to health systems research, as the importance of context is large, there are also likely to be multiple interpretations and perspectives that could be important to capture, and it is amenable to examining the relationships between actors.
3. The ethnographic lens: adaptations of ethnographic methods such as participant observation focusing on groups of practitioners in their socialised professional settings and organisations.

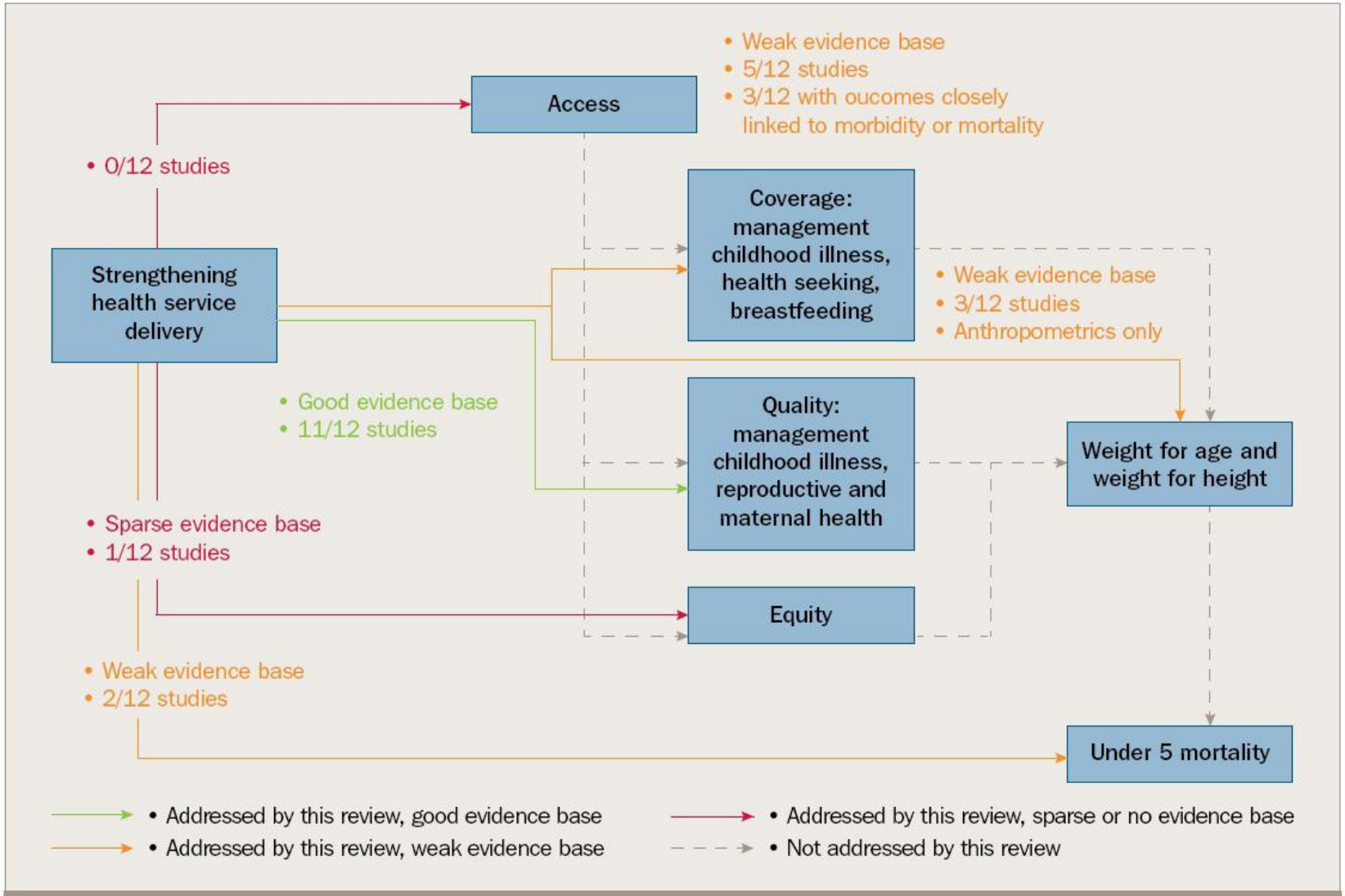
4. Advances in impact evaluation: focusing on impact as a causal mechanism, rather than an epidemiological definition of impact, as for example, mortality. Moving beyond experimental designs to other methodologies for establishing a counterfactual, with an emphasis on external validity/generalisability.

5. Investigating policy and system change over time: including longitudinal designs with repeated measures, but also historical research methods and policy analysis.

6. Cross-national analysis: due to worries about generalisability and applicability, cross-national analyses have been suggested, although due to scale, complexity and cost these are relatively uncommon.

7. Action research: principally a qualitative interactive inquiry process implemented in a collaborative context with data-driven analysis.

Figure 5.1: Potential pathways of effect and mapping the evidence base



Mapping the availability of evidence to the logic framework: linking strengthening health service delivery and mortality

5.4 Shared characteristics of success

Findings from synthesising the effectiveness evidence show that of the eleven studies that reported quality outcomes, eight showed substantial and consistent improvement following the intervention.

All seven studies classified in categories B and C, were successful in terms of improving quality, while only one of the four studies in category A that measured quality was successful. This suggests that interventions that moved beyond access to information and training alone (i.e. improving technical guidance) to incorporate activities to strengthen other components of health service delivery or other levels at which constraints may act showed improved results in terms of quality of care.

This finding is particularly apparent for outcomes of quality of health care. However, it can only loosely be extended to outcomes of coverage, and evidence for the effectiveness of included interventions on survival impact is currently statistically inconclusive, although non-significant trends were reported.

Similar evidence to suggest that wider elements of health service delivery at the health policy and strategic management level are important in delivering health services effectively at scale has recently been reported by Goga and colleagues in their study of challenges with scaling up IMCI, which principally drew on reporting from key informants from 27 LMICs implementing IMCI (Goga and Muhe 2011).

Within these eight studies, we identified some shared characteristics relating to the interventions, based on the published descriptions. These have been broadly summarised into four characteristics, and Table 5.1 shows the distribution of these characteristics across the eight studies.

The shared characteristics

1. **Sustained interaction**: studies with repeated interactions between the intervention and the health worker, and those that had continued interaction over a substantial period of time tended to show more consistently positive findings.

2. **Local ownership through a focal point person**: many of the studies reporting successful outcomes incorporated an individual, often an existing health facility staff member, to act as the on-site point person for the intervention, which seemed to be important in fostering ownership of the intervention by participating facilities.

3. **Feedback cycles**: studies that incorporated feedback to health staff, often through continued supportive supervision or problem solving, although sometimes through interim monitoring and assessments, showed more consistently positive findings.

4. **Beyond the health facility**: many of the studies reporting successful quality outcomes, included aspects of the intervention that went beyond clinical management and targeted management, supervision and wider systems management (e.g. record keeping, supplies monitoring) at the health facility level, but also linked these processes and their findings to management and planning systems beyond the health facility, for example to the district.

Table 5.1: Shared characteristics of success

Intervention category	**(A) Strengthening weak technical guidance only**	**(B) Strengthening weak technical guidance and other elements of the service delivery level**				**(C) Strengthening weak technical guidance and other elements of the service delivery level or other levels on which there are constraints to improving access to health care**		
Study and reference	**Kenya (Zurovac et al., 2011) SMS malaria**	**Benin (Jennings et al., 2010) ANC job aids:**	**Benin (Rowe et al., 2009) Post IMCI supervision**	**Guinea and Kenya (Bradley and Igras, 2005) Health centre quality improvement**	**Kenya (Ayieko et al., 2011) Implementation of guidelines**	**Mali (Winch et al., 2003) Community treatment of malaria**	**Tanzania (Armstrong Schellenberg et al., 2004a, b) IMCI**	**Bangladesh (Arifeen et al., 2004a, b) IMCI**
Sustained interaction	✓		✓	✓	✓			
Local ownership through a focal point person		✓	✓	✓	✓			
Feedback cycles		✓	✓	✓	✓		✓	✓
Beyond the health facility				✓	✓	✓	✓	✓

5.5 Limitations to this review

5.5.1 Search strategy

Constructing search strategies and choosing appropriate search terms for health systems research is considerably less straightforward than for clinical interventions. There is also a balance to be achieved between sensitivity and specificity; although it is essential to maximise exposure of reviewers screening citations to any studies that are potentially eligible, this must be balanced with feasibility. The number of citations is comparable with many of the other systematic reviews in this area that were summarised above. An exception is the wide-ranging, comprehensive review underway by Rowe and colleagues, who have searched almost 45 electronic sources and screened over 105,000 citations, albeit over a period of six years. As such, we do not think that the search strategy employed in this review is a serious limitation.

5.5.2 Grey literature

Considerable in-country effort is ongoing to improve health service delivery in LMICs, often occurring in partnership between ministries of health and numerous multi-, bi-lateral and funding organisations, including but not limited to The World Bank, USAID, The Global Alliance for Vaccines and Immunization (GAVI), The Global Fund to Fight AIDS, Tuberculosis and Malaria, The UK Department for International Development (DfID), and Management Services for Health (MSH) (Vujicic et al., 2012). This review highlights that the accessible evaluations to address questions of effectiveness of approaches to strengthen health service delivery implemented at scale in LMICs using robust designs are considerably fewer.

However the search strategy devised did not perform well in identifying grey literature, which could introduce bias. This may have an impact on the completeness of the findings of this review in relation to the first objective of describing and characterising interventions to improve health service delivery, as the type of evaluation excluded from this review, particularly that which may be found more prominently in the grey literature, may provide additional and different categories of interventions to improve health service delivery. As such, additional learning for objective one of this review may be available using different review criteria. For a systematic review focusing on effectiveness, this may be less problematic, but could be very important when considering questions of implementation, examining barriers to and facilitators of implementation, and applicability to other contexts.

5.5.3 Inclusion criteria: study design

This review identified few studies meeting the inclusion criteria, although the numbers are not dissimilar to other systematic reviews that focused on robust study designs in LMICs. As evidenced by other systematic reviews summarised, restricting the focus to LMICs reduces the evidence base available. However, the most substantial effect on number of studies included is from restrictions in study design. This was necessary to address questions of effectiveness, but nonetheless highlights that there are still improvements to be made to the rigour with which health systems interventions in LMICs are evaluated.

In terms of addressing questions of effectiveness, controlled studies provide high internal validity, and may answer questions of whether an intervention worked under specified circumstances. However, randomised methodologies, even cluster studies, are frequently not the most feasible, practical or desirable study design to evaluate interventions delivered at scale. They may also have low external validity

(generalisability), particularly for the evaluation of complex interventions. Therefore, in considering wider health systems research and policy implementation questions, which may include questions of the operational success of interventions, assessment of how and why an intervention worked, identifying whether the intervention will work in other settings, or allowing clear lessons across studies to be identified, additional study designs are required, as well as detailed information on context (Victora et al., 2011).

5.5.4 Inclusion criteria: scale

This review was unique in restricting inclusion criteria to studies carried out 'at scale', defined as in at least one district. Three papers which underwent full-paper review were excluded due to insufficient scale (Bhunia et al., 2010, Darmstadt et al., 2010, Hermida and Robalino, 2002); in addition, another three studies were excluded at title and abstract screening stage due to insufficient scale (Bhutta et al., 2008, Brown et al., 2008, Penny et al., 2005). Dr Alexander Rowe kindly provided us with a copy of the included public health sector (i.e. state delivery) studies included in his much more wide-ranging review (covering interventions to address constraints at the household, community, health service delivery and health sector policy and strategic management levels). By examining this list, we identified a further three studies that would have met our review's inclusion criteria except for that of scale, but which were not found using our search strategy (Biai et al., 2007, Johnson et al., 2002, Lueveswanij et al., 2000). Therefore, although removing the restriction of scale would have included at least nine additional studies, this could still not be said to result in a review that would constitute a wide evidence base.

Furthermore, 'at scale' is not a precisely defined term, and in this review, a pragmatic definition of at scale was used, referring to implementation at the district level. In practice, experience from only one or a few districts may not be sufficient to elucidate questions beyond efficacy and effectiveness.

5.5.5 Time period, language, and analysis

Some relevant evidence may have been missed due to restricting the time frame of this review to studies published from 2000 onwards; however, given the focus of this review on interventions delivered at scale, this time period was informed by the increased interest in scaling up in the past decade (Mangham and Hanson, 2010).

Finally, although no language restrictions were applied to the search strategy, the search terms used were in English, and we observed that few Spanish-language citations were retrieved. It is likely therefore that relevant evidence from Latin America is missing, and it is difficult to conclude how applicable findings from this review are to low and lower middle income countries outside Africa and Asia.

We also highlight some limitations to our analysis. Due to the broad range of included interventions and outcomes, meta-analysis was not appropriate. Data were aggregated instead based on narrative synthesis using a thematic summary approach (Thomas et al., 2012). The broad nature of the systematic review question and the focus on the effectiveness of interventions did not enable detailed logic frameworks for each intervention that identified and illustrated interim processes and outputs between intervention inputs and assessed outcomes. The thematic summary approach is susceptible to pitfalls of 'vote counting'; however, we have tried to guard against this in our overall synthesis and

maintained a focus on the quality of studies, rather than simply on the number of studies (Waddington et al., 2012).

5.6 Use of systematic reviews in health systems research

In addition to the limitations of this review, it is important to reflect on the role of systematic reviews within health systems research more generally.

Systematic reviews and meta- analyses were developed during the 1970s to examine relatively simple, often pharmaceutical interventions and their effect on patient clinical outcomes. It remains that the most simply synthesised evidence is obtained from reviewing simple interventions. However, important questions in health often relate to more complex interventions, and these are of interest to policy makers.

The past five years or so have seen the application of systematic review methodology to health systems and development interventions relevant to policy making (White and Waddington, 2012). They suggest three areas to which systematic reviews of such interventions may contribute:

1. defining and framing a problem;
2. assessing potential policy options;
3. identifying implementation considerations for selected policy options.

They highlight that crucial within this is specifying the role a systematic review will fulfil for policy-makers and other stakeholders. It is possible then for methodology experts to work with policy makers to determine an appropriate strategy to attempt to fulfil this role. Snilstveit and colleagues highlight that a tension can exist between the broad questions that a policy maker needs answered in a timely fashion to inform and contribute to important decision-making processes, and defining a question amenable to systematic review methodology (Snilstveit, 2012).

Adaptations to traditional systematic review methodology can be made to fit the three purposes outlined. For example, scoping reviews using robust methodology but not adhering to strict systematic review criteria, or rapid reviews of existing systematic reviews, may be a helpful and timely way of defining and framing a problem, and identifying what is already known about an area (Stewart et al., 2012).

Systematic reviews themselves may also contribute to assessing potential policy options by serving as a type of ‘audit’ of the evidence base (White and Waddington, 2012). However, within this, it is important to identify which evidence base is of use to policy makers (Petticrew and Roberts, 2006):

- Is this the efficacy and effectiveness evidence base of a relatively narrowly defined intervention whose impact or effect on a narrowly defined outcome or small set of outcomes can be numerically summarised with meta-analysis?
    - This scenario is close to the traditional pharmaceutical role of systematic reviews and meta-analysis.
- Or is this the evidence base of a wide set of interventions on a number of outcomes?

- In this case, it is important to relax study design criteria in order to avoid 'empty' reviews. Criteria can still be set to determine the types of studies of appropriate rigour to answer the question asked. There may be challenges in systematically assessing the quality of these study designs- although some guidance exists (Higgins and Green, 2011, Waddington et al., 2012).

These types of systematic review may be useful in highlighting gaps in the evidence base. However, in this scenario, the importance of identifying literature from grey sources and databases outside the well-known and frequently used medical databases is even more important. As is using search strategies that are appropriate to the sources searched; indexing applied within medical databases is not standard across other data sources (Snilstveit, 2012).

It is feasible within efficacy or effectiveness reviews of a narrowly defined intervention to extend the systematic review to what Snilstveit and colleagues term 'effectiveness plus', by additionally identifying, extracting and synthesising evidence on not only 'what' works, but expanding the review to questions of 'how', why', 'under what circumstances' and 'for whom' (Snilstveit, 2012). This can expand the role of systematic reviews to the third of White's categories - identifying implementation considerations for selected policy options (White and Waddington, 2012).

However, this requires appropriate searching and locating of relevant evidence, which may be a different strategy to that for identifying the effectiveness evidence. It probably involves the inclusion of qualitative data and information from sources (e.g. protocols) that may be hard to access, and contact with authors and researchers directly. Methods for assessing the quality of these study designs and data are less well developed in some cases, and generally less well known.

In addition, Snilstveit and colleagues recommend the incorporation of theory into such systematic reviews by the explicit inclusion of a logic model or causal pathway analysis, and mapping the evidence along this (for example, Stewart et al., 2012). However, this requires a relatively narrow focus in terms of intervention and outcomes.

Snilstveit and colleagues highlight that for these types of 'effectiveness plus' systematic reviews, that reporting of the interventions must be detailed, information on the implementation process and fidelity of implementation must be documented, and systematic and detailed documentation of context is necessary in order for policy makers and other stakeholders across a variety of settings to draw out lessons of applicability (whether the intervention, as implemented, can be implemented in a new setting) and transferability (whether the study findings can be replicated in a new setting) (Snilstveit, 2012).

This review sits mostly within the second area of contribution identified by White and colleagues, of assessing potential policy options, and specifically assessing those where evidence for effectiveness from at-scale implementation in low and lower middle income countries is available. The review question challengingly sat between an emphasis on evidence for effectiveness, but without being restricted to a narrowly defined set of interventions and one or few primary outcomes, with quite a range of interventions and outcomes targeted. This review has highlighted that the evidence base for outcomes of quality is largest, but that the evidence base for other outcomes, including survival impact, is scarce in the published literature. Interventions were too diverse to summarise using meta-analysis;

however, effectiveness was summarised using narrative synthesis and the interventions characterised using a framework to enable comparisons across interventions. Interventions were too diverse to map out a detailed logic model or interim steps within a causal pathway. This review does not extend to fulfil the role of an 'effectiveness plus' systematic review, but efforts have been made to draw attention to shared characteristics of success across the 12 included studies, to provide some lessons in applicability and transferability.

6. Implications

We anticipate that users of this research will include donor agencies, academics and others, and acknowledge that they will want to undertake a process of interpretation and application of the results of this review. However, on the basis of our findings we draw out implications for policy makers and for research in Table 6.1.

Table 6.1: Implications for policy makers and for research

Observation from this and/or other systematic reviews	Implications for policy makers	Implications for research
Few studies assess survival impact, with outcomes of health worker knowledge and behaviour more common	Knowledge on impact remains low. There is some evidence on under-five and neonatal mortality from fully-implemented IMCI (i.e. not only the training components) from this review, and of using lay health workers from other systematic reviews Policy makers should determine the desired and necessary level of evidence, and whether evidence of impact is necessary	If knowledge of impact is required, it needs to be measured, or things that will allow us to model it need to be measured, e.g. using the Lives Saved tool (LiST)[4] Logic frameworks are useful for helping to think through and test causal pathways
Technical guidance alone is insufficient	Interventions to improve health worker knowledge and behaviour should move beyond didactic training	If it is useful to policy makers to identify what additional components are essential/ sufficient to improve behaviour, alternative strategies must be compared
Quality of the evidence	The quality of the evidence available from studies in LMICs, based on GRADE criteria, is generally low It is necessary to determine the appropriate desired level of quality to inform policy	If it is appropriate, studies designed to generate moderate- to high-quality evidence should be undertaken in LMICs These studies should also take into consideration the points below
Why, how, for whom, under what circumstances	This type of information will help address questions of applicability and transferability to other	Studies should incorporate careful documentation and reporting of implementation, including fidelity and

[4] http://www.jhsph.edu/departments/international-health/IIP/list/

Observation from this and/or other systematic reviews	Implications for policy makers	Implications for research
	settings than that of the original study	adaptations Careful reporting of context, barriers and facilitators to implementation is needed Qualitative methods should be incorporated as core elements of the design and evaluation of complex interventions
The use of theory	Theory and logic models of causal pathways are not frequently used These may help to address questions of applicability and transferability to other settings than that of the original study	The incorporation of theory and logic models should be encouraged A holistic approach to designing and evaluating health system interventions may be useful Studies should also take into account the next point
Evidence is required on sustainability, cost-effectiveness, acceptability, equity	Evidence in these other key aspects that are relevant to decision making are lacking	In addition to a focus on outcomes beyond health worker knowledge and behaviour, studies should consider including these outcomes

7. References

Adam T, Hsu J, de Savigny D, Lavis JN, Rottingen JA, Bennett S (2012) Evaluating health systems strengthening interventions in low-income and middle-income countries: are we asking the right questions? *Health Policy and Planning* 27 (Suppl. 4): 9-19.

Akter SFU, Heller RD, Smith AJ, Milly AF (2009) Impact of a training intervention on use of antimicrobials in teaching hospitals. *Journal of Infection in Developing Countries* 3 (6): 447-451.

Anderson LM, Petticrew M, Rehfuess E, Armstrong R, Ueffing E, Baker P, Francis D, Tugwell P (2011) Using logic models to capture complexity in systematic reviews. *Research Synthesis Methods* 2 (1): 33-42.

Arifeen SE, Blum LS, Emdadul Hoque DM, Chowdhury EK, Khan R, Black PRE, Victora PCG, Bryce J (2004a) Integrated management of childhood illness (IMCI) in Bangladesh: early findings from a cluster-randomised study. *Lancet* 364 (9445): 1595-1602.

Arifeen SE, Hoque DME, Akter T, Rahman M, Hoque ME, Begum K, Chowdhury EK, Khan R, Blum LS, Ahmed S, Hossain MA, Siddik A, Begum N, Sadeq-ur Rahman Q, Haque TM, Billah SM, Islam M, Rumi RA, Law E, Al-Helal ZAM, Baqui AH, Schellenberg J, Adam T, Moulton LH, Habicht J-P, Scherpbier RW, Victora CG, Bryce J, Black RE (2004b) Effect of the integrated management of childhood illness strategy on childhood mortality and nutrition in a rural area in Bangladesh: a cluster randomised trial. *Lancet* 374 (9687): 393-403.

Armstrong Schellenberg J, Bryce J, de Savigny D, Lambrechts T, Mbuya C, Mgalula L, Wilczynska K (2004a) The effect of integrated management of childhood illness on observed quality of care of under-fives in rural Tanzania. *Health Policy and Planning* 19 (1): 1-10.

Armstrong Schellenberg JR, Adam T, Mshinda H, Masanja H, Kabadi G, Mukasa O, John T, Charles S, Nathan R, Wilczynska K, Mgalula L, Mbuya C, Mswia R, Manzi F, de Savigny D, Schellenberg D, Victora C (2004b) Effectiveness and cost of facility-based integrated management of childhood illness (IMCI) in Tanzania. *Lancet* 364 (9445): 1583-1594.

Ayieko P, Ntoburi S, Wagai J, Opondo C, Opiyo N, Migiro S, Wamae A, Mogoa W, Were F, Wasunna A, Fegan G, Irimu G, English M (2011) A multifaceted intervention to implement guidelines and improve admission paediatric care in Kenyan district hospitals: a cluster randomised trial. *PLoS Medicine* 8: 4.

Bennett S, Adam T, Zarowsky C, Tangcharoensathien V, Ranson K, Evans T, Mills A (2008) From Mexico to Mali: progress in health policy and systems research. *Lancet* 372 (9649): 1571-1578.

Bhunia R, Hutin Y, Ramkrishnan R, Ghosh PK, Dey S, Murhekar M (2010) Reducing use of injections through interactional group discussions: a randomized controlled trial. *Indian Pediatrics* 47 (5): 409-414.

Bhutta ZA, Memon ZA, Soofi S, Salat MS, Cousens S, Martines J (2008) Implementing community-based perinatal care: results from a pilot study in rural Pakistan. *Bulletin of the World Health Organization* 86 (6): 452-459.

Bhutta ZA, Soofi S, Cousens S, Mohammad S, Memon ZA, Ali I, Feroze A, Raza F, Khan A, Wall S, Martines J (2011) Improvement of perinatal and newborn care in rural Pakistan through community-based strategies: a cluster-randomised effectiveness trial. *Lancet* 377(9763): 403-412.

Biai S, Rodrigues A, Gomes M, Ribeiro I, Sodemann M, Alves F, Aaby P (2007) Reduced in-hospital mortality after improved management of children under 5 years admitted to hospital with malaria: randomised trial. *British Medical Journal* 335 (7625): 862.

Bosch-Capblanch X, Liaqat S, Garner P (2011) Managerial supervision to improve primary health care in low- and middle-income countries. *Cochrane Database of Systematic Reviews: Reviews 2011*. Issue 9, DOI: 10.1002/14651858.CD006413.pub2.

Bradley J, Igras S (2005) Improving the quality of child health services: participatory action by providers. *International Journal for Quality in Health Care* 17 (5): 391-399.

Briggs CJ, Garner P (2006) Strategies for integrating primary health services in middle- and low-income countries at the point of delivery. *Cochrane Database of Systematic Reviews* 2: CD003318.

Briggs CJ, Capdegelle P, Garner P (2001) Strategies for integrating primary health services in middle- and low-income countries: effects on performance, costs and patient outcomes. *Cochrane Database of Systematic Reviews* 4: CD003318.

Brown LD, de Negri B, Hernandez O, Dominguez L, Sanchack JH, Roter D (2008) An evaluation of the impact of training Honduran health care providers in interpersonal communication. *International Journal for Quality in Health Care* 12 (6): 495-501.

Bryce J, el Arifeen S, Pariyo G, Lanata CF, Gwatkin D, Habicht J-P (2003) Reducing child mortality: can public health deliver? *The Lancet* 362 (9378): 159-164.

Bryce J, Gouws E, Adam T, Black RE, Schellenberg JA, Manzi F, Victora CG, Habicht J-P (2005) Improving quality and efficiency of facility-based child health care through integrated management of childhood illness in Tanzania. *Health Policy and Planning* 20 (Suppl. 1): 69-76.

Carlo WA, Goudar SS, Jehan I, Chomba E, Garces A, Parida S, Althabe F, McClure EM, Derman RJ, Goldenberg RL, Bose C, Krebs NF, Panigrahi P, Buekens P, Chakraborty H, Hartwell TD, Wright LL, Belizan JM, Ditekemena J, Tshefu A, Mohapatra A, Kodkany BS, Jafri SR, Pasha O, Mbelenga M, Hambidge M, Douoguih M, Willoughby A, Harris H (2010) Newborn-care training and perinatal mortality in developing countries. *New England Journal of Medicine* 362 (7): 614-623.

Claeson M, Gillespie D, Mshinda H, Troedsson H, Victora CG (2003) Knowledge into action for child survival. *Lancet* 362 (9380): 323-327.

Craig PDP, Macintyre S, Michie S, Nazareth I, Petticrew M (2008) *Developing and evaluating complex interventions: new guidance*. London: Medical Research Council. www.mrc.ac.uk/complexinterventionsguidance (accessed 14 March 2013).

Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, de Bernis L (2005) Evidence-based, cost-effective interventions: how many newborn babies can we save? *The Lancet* 365 (9463): 977-988.

Darmstadt GL, Choi Y, Arifeen SE, Bari S, Rahman SM, Mannan I, Seraji HR, Winch PJ, Saha SK, Ahmed ASMNU, Ahmed S, Begum N, Lee ACC, Black RE, Santosham M, Crook D, Baqui AH, Bangladesh Projahnmo-2 Mirzapur Study Group (2010) Evaluation of a cluster-randomized controlled trial of a package of community-based maternal and newborn interventions in Mirzapur, Bangladesh. *PLoS ONE* 5: 3. http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0009696 (accessed 14 March 2013).

Davidoff F, Batalden P, Stevens D, Ogrinc G, Mooney SE (2009) Publication guidelines for quality improvement studies in health care: evolution of the SQUIRE project. *British Medical Journal* 338: a3512.

de Savigny D, Adam T (2009) *Systems thinking for health systems strengthening*. Geneva: World Health Organization.

Donabedian A (2003) *An introduction to quality assurance in health care*. Oxford: Oxford University Press.

Dudley L, Garner P (2011) Strategies for integrating primary health services in low- and middle-income countries at the point of delivery. *Cochrane Database of Systematic Reviews* 7: CD003318.

Eldridge C, Palmer N (2009) Performance-based payment: some reflections on the discourse, evidence and unanswered questions. *Health Policy and Planning* 24 (3): 160-166.

Forsetlund L, Bjorndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf F, Davis D, Odgaard-Jensen J, Oxman AD (2009) Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 2: CD003030.

Frenk J (2010) The global health system: strengthening national health systems as the next step for global progress, *PLoS Medicine* 7 (1): e1000089. http://dx.doi.org/10.1371%2Fjournal.pmed.1000089 (accessed 14 March 2013).

Fryatt R, Mills A, Nordstrom A (2010) Financing of health systems to achieve the health Millennium Development Goals in low-income countries. *The Lancet* 375 (9712): 419-426.

Gilroy K, Winch PJ, Diawara A, Swedberg E, Thiero F, Kane M, Daou Z, Berthe Z, Bagayoko A (2004) Impact of IMCI training and language used by provider on quality of counseling provided to parents of sick children in Bougouni District, Mali. *Patient Education and Counseling* 54 (1): 35-44.

Gilson L (ed.) (2012) *Health policy and systems research: a methodology reader*. Geneva: World Health Organization. http://www.who.int/alliance-hpsr/alliancehpsr_reader.pdf (accessed 14 March 2013).

Goga A, Muhe L (2011) Global challenges with scale-up of the integrated management of childhood illness strategy: results of a multi-country survey. *BMC Public Health* 11 (1): 503. http://www.biomedcentral.com/1471-2458/11/503 (accessed 14 March 2013).

Grobler L, Marais BJ, Mabunda SA, Marindi PN, Reuter H, Volmink J (2009) Interventions for increasing the proportion of health professionals practising in rural and other underserved areas. *Cochrane Database of Systematic Reviews* 1: CD005314. doi: 10.1002/14651858.CD005314.pub2.

Gulmezoglu AM, Langer A, Piaggio G, Lumbiganon P, Villar J, Grimshaw J (2007) Cluster randomised trial of an active, multifaceted educational intervention based on the WHO Reproductive Health Library to improve obstetric practices. *BJOG: An International Journal of Obstetrics and Gynaecology* 114 (1): 16-23.

Hanson K, Ranson K, Oliveira-Cruz V, Mills A (2003) Expanding access to priority health interventions: a framework for understanding the constraints to scaling-up. *Journal of International Development* 15: 1-14.

Hermida J, Robalino ME (2002) Increasing compliance with maternal and child care quality standards in Ecuador. *International Journal for Quality in Health Care* 14 (Suppl. 1): 25-34.

Higgins JPT, Green S (eds) (2011) *Cochrane handbook for systematic reviews of interventions*. http://handbook.cochrane.org/ (accessed 14 March 2013).

Huicho L, Dieleman M, Campbell J, Codjia L, Balabanova D, Dussault G, Dolea C (2010) Increasing access to health workers in underserved areas: a conceptual framework for measuring results. *Bulletin of the World Health Organization* 88 (5): 357-363.

Jamtvedt G, Young JM, Kristoffersen DT, Thomson O'Brien MA, Oxman AD (2003) Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 3: CD000259.

Jennings L, Yebadokpo AS, Affo J, Agbogbe M (2010) Antenatal counseling in maternal and newborn care: use of job aids to improve health worker performance and maternal understanding in Benin. *BMC Pregnancy Childbirth* 10: 75. http://www.biomedcentral.com/1471-2393/10/75 (accessed 14 March 2013).

Johnson KW, Young LC, Suresh G, Berbaum ML (2002) Drug abuse treatment training in Peru: a social policy experiment. *Evaluation Review* 26 (5): 480-519.

Jokhio AH, Winter HR, Cheng KK (2005) An intervention involving traditional birth attendants and perinatal and maternal mortality in Pakistan. *New England Journal of Medicine* 352 (20): 2091-2099.

Jones G, Steketee RW, Black RE, Bhutta ZA, Morris SS (2003) How many child deaths can we prevent this year? *The Lancet* 362 (9377): 65-71.

Kelley E, Geslin C, Djibrina S, Boucar M (2001) Improving performance with clinical standards: the impact of feedback on compliance with the integrated management of childhood illness algorithm in Niger, West Africa. *International Journal of Health Planning and Management* 16 (3): 195-205.

Kumar V, Mohanty S, Kumar A, Misra RP, Santosham M, Awasthi S, Baqui AH, Singh P, Singh V, Ahuja RC, Singh JV, Malik GK, Ahmed S, Black RE, Bhandari M, Darmstadt GL, Saksham Study Group (2008) Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India: a cluster-randomised controlled trial. *Lancet* 372 (9644): 1151-1162.

Lagarde M, Palmer N (2009) The impact of contracting out on health outcomes and use of health services in low and middle-income countries. *Cochrane Database of Systematic Reviews* 4: CD008133.

Lagarde M, Palmer N (2011) The impact of user fees on access to health services in low- and middle-income countries. *Cochrane Database of Systematic Reviews* 4: CD009094.

Lagarde M, Haines A, Palmer N (2007) Conditional cash transfers for improving uptake of health interventions in low- and middle-income countries: a systematic review. *Journal of the American Medical Association* 298 (16): 1900-1910.

Lagarde M, Haines A, Palmer N (2009) The impact of conditional cash transfers on health outcomes and use of health services in low and middle income countries. *Cochrane Database of Systematic Reviews* 4: CD008137.

Lancet (2010) Evaluation: the top priority for global health. *The Lancet* 375 (9714): 526.

Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, Jack W, Habyarimana J, Sadatsafavi M, Najafzadeh M, Marra CA, Estambale B, Ngugi E, Ball TB, Thabane L, Gelmon LJ, Kimani J, Ackers M, Plummer FA (2010) Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (Weltel Kenya1): a randomised trial. *Lancet* 376 (9755): 1838-1845.

Lewin S, Munabi-Babigumira S, Glenton C, Daniels K, Bosch-Capblanch X, van Wyk BE, Odgaard-Jensen J, Johansen M, Aja GN, Zwarenstein M, Scheel IB (2010) Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases. *Cochrane Database of Systematic Reviews* 3: CD004015. doi: 10.1002/14651858.CD004015.pub3.

Lueveswanij S, Nittayananta W, Robison VA (2000) Changing knowledge, attitudes, and practices of Thai oral health personnel with regard to AIDS: an evaluation of an educational intervention. *Community Dental Health* 17 (3): 165-171.

Madon T, Hofman KJ, Kupfer L, Glass RI (2007) Implementation science. *Science* 318 (5857): 1728-1729. http://www.sciencemag.org/content/318/5857/1728.short (accessed 14 March 2013).

Mangham LJ, Hanson K (2010) Scaling up in international health: what are the key issues? *Health Policy and Planning* 25 (2): 85-96.

Masanja H, Schellenberg JA, de Savigny D, Mshinda H, Victora CG (2005) Impact of integrated management of childhood illness on inequalities in child health in rural Tanzania. *Health Policy and Planning* 20 (Suppl. 1): i77-i84. http://heapol.oxfordjournals.org/content/20/suppl_1/i77.abstract (accessed 14 March 2013).

Matendo R, Engmann C, Ditekemena J, Gado J, Tshefu A, Kinoshita R, McClure EM, Moore J, Wallace D, Carlo WA, Wright LL, Bose C (2011) Reduced perinatal mortality following enhanced training of birth attendants in the Democratic Republic of Congo: a time-dependent effect. *BMC Medicine* 9: 93.

Mbeba MM, Kaponda CPN, Jere DL, Kachingwe SI, Crittenden KS, McCreary LL, Norr JL, Norr KF (2011) Peer group intervention reduces personal HIV risk for Malawian health workers. *Journal of Nursing Scholarship* 43 (1): 72-81.

McIntyre D, Thiede M, Birch S (2009) Access as a policy-relevant concept in low- and middle-income countries. *Health Economics, Policy and Law* 4 (2): 179-193.

Murray CJ, Frenk J (2000) A framework for assessing the performance of health systems. *Bulletin of the World Health Organization* 78 (6): 717-731.

Naimoli JF, Rowe et al., Lyaghfouri A, Larbi R, Lamrani LA (2006) Effect of the integrated management of childhood illness strategy on health care quality in Morocco. *International Journal for Quality in Health Care* 18 (2): 134-144.

Ngasala B, Mubi M, Warsame M, Petzold MG, Massele AY, Gustafsson LL, Tomson G, Premji Z, Bjorkman A (2008) Impact of training in clinical and microscopy diagnosis of childhood malaria on antimalarial drug prescription and health outcome at primary health care level in Tanzania: a randomized controlled trial. *Malaria Journal* 7: 199.

Oakley A, Strange V, Bonell C, Allen E, Stephenson J (2006) Process evaluation in randomised controlled trials of complex interventions. *British Medical Journal* 332: 413-416.

Oliveira-Cruz V, Hanson K, Mills A (2003) Approaches to overcoming constraints to effective health service delivery: a review of the evidence. *Journal of International Development* 15 (1): 41-65.

Opiyo N, English M (2010) In-service training for health professionals to improve care of the seriously ill newborn or child in low and middle-income countries (review). *Cochrane Database of Systematic Reviews* 4: CD007071. doi: 10.1002/14651858.CD007071.pub2.

Pagaiya N, Garner P (2005) Primary care nurses using guidelines in Thailand: a randomized controlled trial. *Tropical Medicine and International Health* 10 (5): 471-477.

Penaloza B, Pantoja T, Bastias G, Herrera C, Rada G (2011) Interventions to reduce emigration of health care professionals from low- and middle-income countries. *Cochrane Database of Systematic Reviews* 9: CD007673. doi: 10.1002/14651858.CD007673.pub2.

Penny ME, Creed-Kanashiro HM, Robert RC, Narro MR, Caulfield LE, Black RE (2005) Effectiveness of an educational intervention delivered through the health services to improve nutrition in young children: a cluster-randomised controlled trial. *Lancet* 365 (9474): 1863-1872.

Peters DH, Kohli M, Mascarenhas M, Rao K (2006) Can computers improve patient care by primary health care workers in India? *International Journal for Quality in Health Care* 18 (6): 437-445.

Petticrew M, Roberts H (2006) *Systematic reviews in the social sciences: a practical guide*. Oxford: Blackwell.

Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, Britten N, Roen K, Duffy S (2006) *Guidance on the conduct of narrative synthesis in systematic reviews*. http://www.lancs.ac.uk/shm/research/nssr/research/dissemination/publications/NS_Synthesis_Guidance_v1.pdf (accessed 14 March 2013).

Remme JHF, Adam T, Becerra-Posada F, D'Arcangues C, Devlin M, Gardner C, Ghaffar A, Hombach J, Kengeya JFK, Mbewu A, Mbizvo MT, Mirza Z, Pang T, Ridley RG, Zicker F, Terry RF (2010) Defining research to improve health systems. *PLoS Medicine* 7 (11): e1001000. http://dx.doi.org/10.1371%2Fjournal.pmed.1001000 (accessed 14 March 2013).

Rowe AK, Onikpo F, Lama M, Osterholt DM, Rowe SY, Deming MS (2009) A multifaceted intervention to improve health worker adherence to integrated management of childhood illness guidelines in Benin. *American Journal of Public Health* 99 (5): 837-846.

Rowe AK, Rowe S, Peters DH, Holloway KA, Chalker J, Ross-Degnan D (2012) Preliminary results of a systematic review of the effectiveness and costs of strategies to improve health worker performance in low- and middle-income countries (LMICS). Paper presented at: *Second Global Symposium on Health Systems Research*, Beijing, China.

Shah SA, Kristensen S, Memon MA, White HL, Vermund SH (2007) Syndromic management training for non-formal care providers in Pakistan improves quality of care for sexually transmitted diseases STD care: a randomized clinical trial. *Southeast Asian Journal of Tropical Medicine and Public Health* 38 (4): 737-748.

Sharma N, Pagare D, Singh MM, Saha R, Ingle GK (2005) Effectiveness of modular training regarding revised national tuberculosis control programme among paramedical workers in Delhi. *Journal of Communicable Diseases* 37 (3): 245-248.

Sheikh K, Gilson L, Agyepong IA, Hanson K, Ssengooba F, Bennett S (2011) Building the field of health policy and systems research: framing the questions. *PLoS Medicine* 8 (8): e1001073. http://dx.doi.org/10.1371%2Fjournal.pmed.1001073

Shepperd S, Lewin S, Straus S, Clarke M, Eccles MP, Fitzpatrick R, Wong G, Sheikh A (2009) Can we systematically review studies that evaluate complex interventions? *PLoS Medicine* 6 (8): e1000086. http://dx.doi.org/10.1371%2Fjournal.pmed.1000086 (accessed 14 March 2013).

Sibley LM, Sipe TA, Brown CM, Diallo MM, McNatt K, Habarta N (2007) Traditional birth attendant training for improving health behaviours and pregnancy outcomes. *Cochrane Database of Systematic Reviews* 3: CD005460.

Skarbinski J, Ouma PO, Causer LM, Kariuki SK, Barnwell JW, Alaii JA, De Oliveira AM, Zurovac D, Larson BA, Snow RW, Rowe et al., Laserson KF, Akhwale WS, Slutsker L, Hamel MJ (2009) Effect of malaria rapid diagnostic tests on the management of uncomplicated malaria with artemether-lumefantrine in Kenya: a cluster randomized trial. *American Journal of Tropical Medicine and Hygiene* 80 (6): 919-926.

Smith LA, Jones C, Meek S, Webster J (2009) Review: provider practice and user behavior interventions to improve prompt and effective treatment of malaria: do we know what works? *The American Journal of Tropical Medicine and Hygiene* 80 (3): 326-335.

Snilstveit B (2012) Systematic reviews: from 'bare bones' reviews to policy relevance. *Journal of Development Effectiveness* 4 (3): 388-408.

Spaan E, Mathijssen J, Tromp N, McBain F, ten Have A, Baltussen R. (2012) The impact of health insurance in Africa and Asia: a systematic review. *Bulletin of the World Health Organization* 90: 685-692.

Stewart R, van Rooyen C, de Wet T (2012) Purity or pragmatism? Reflecting on the use of systematic review methodology in development. *Journal of Development Effectiveness* 4 (3): 430-444.

Thomas J, Harden A, Newman M (2012) Synthesis: combining results systematically and appropriately. In: Gough D, Oliver S and Thomas J (eds) *An introduction to systematic reviews.* London: Sage, pages 179-226.

Travis P, Bennett S, Haines A, Pang T, Bhutta Z, Hyder AA, Pielemeier NR, Mills A, Evans T (2004) Overcoming health-systems constraints to achieve the Millennium Development Goals. *The Lancet* 364 (9437): 900-906.

UNICEF (2012) *Levels and trends in child mortality: report 2012.* New York: UNICEF.

Van Belle SB, 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): 1-10.

van Lonkhuijzen L, Dijkman A, van Roosmalen J, Zeeman G, Scherpbier A (2010) A systematic review of the effectiveness of training in emergency obstetric care in low-resource environments. *BJOG: An International Journal of Obstetrics and Gynaecology* 117 (7): 777-787.

van Olmen J, Criel B, Van Damme W, Marchal B, Van Belle S, Van Dormael M, Hoerée T, Pirard M, Kegels G (2010) *Analysing health systems to make them stronger*. Antwerp: ITGPress.

van Olmen J, Criel B, Bhojani U, Marchal B, Van Belle S, Chenge M, Hoerée T, Pirard M, Van Damme W, Kegels G (2012a) The health system dynamics framework: the introduction of an analytical model for health system analysis and its application to two case-studies. *Health Culture and Society* 2 (1): 2-21.

van Olmen J, Criel B, Van Damme W, Marchal B, Van Belle S, Van Dormael M, Hoerée T, Pirard M, Kegels G (2012b) *Analysing health system dynamics: a framework*. Antwerp: ITGPress.

Victora CG, Black RE, Boerma JT, Bryce J (2011) Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations. *Lancet* 377 (9759): 85-95.

Vujicic, Weber S, Nikolic I, Atun R, Kumar R (2012) An analysis of GAVI, the Global Fund and World Bank support for human resources for health in developing countries. *Health Policy and Planning* 13: 13.

Waddington H, White H, Snilstveit B, Hombrados JG, Vojtkova M, Davies P, Bhavsar A, Eyers J, Koehlmoos TP, Petticrew M, Valentine JC, Tugwell P (2012) How to do a good systematic review of effects in international development: a tool kit. *Journal of Development Effectiveness* 4 (3): 359-387.

White H, Waddington H (2012) Why do we care about evidence synthesis? An introduction to the special issue on systematic reviews. *Journal of Development Effectiveness* 4 (3): 351-358.

Winch PJ, Bagayoko A, Diawara A, Kane M, Thiero F, Gilroy K, Daou Z, Berthe Z, Swedberg E (2003) Increases in correct administration of chloroquine in the home and referral of sick children to health facilities through a community-based intervention in Bougouni District, Mali. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 97 (5): 481-490.

Witter S, Kessy FL, Fretheim A, Lindahl AK (2012) Paying for performance to improve the delivery of health interventions in low and middle-income countries. *Cochrane Database of Systematic Reviews 2*: CD007899. doi: 10.1002/14651858.CD007899.pub2.

World Health Organization (2007) *Everybody business: Strengthening health systems to improve health outcomes: WHO's framework for action.* Geneva: WHO.

World Health Organization (2009) *Global health risks: mortality and burden of disease attributable to selected major risks*. Geneva: WHO.

World Health Organization (2010a) *Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies*. Geneva: WHO.

World Health Organization (2010b) *Countdown to 2015 decade report (2000-2010): taking stock of maternal, newborn and child survival*. Geneva: WHO.

Zaman S, Ashraf RN, Martines J (2008) Training in complementary feeding counselling of healthcare workers and its influence on maternal behaviours and child growth: a cluster-randomized controlled trial in Lahore, Pakistan. *Journal of Health, Population and Nutrition* 26 (2): 210-222.

Zhang T, Wu Y-Q, Zhang X, Xiong Q, Wang Y-P, Zhao G-L, Chen M-H, Ni Z-Z (2004) [an evaluation of effects of intervention on maternal and child health in the rural areas of China] *Sichuan da Xue Xue Bao: Yi Xue Ban/Journal of Sichuan University: Medical Science Edition* 35 (4): 539-542.

Zurovac D, Sudoi RK, Akhwale WS, Ndiritu M, Hamer DH, Rowe et al., Snow RW (2011) The effect of mobile phone text-message reminders on Kenyan health workers' adherence to malaria treatment guidelines: a cluster randomised trial. *Lancet* 378 (9793): 795-803.

Appendices

Appendix 1.1: Authorship of this review

Barbara Willey, London School of Hygiene and Tropical Medicine

Lucy Smith Paintain, London School of Hygiene and Tropical Medicine

Lindsay Mangham-Jefferies, London School of Hygiene and Tropical Medicine

Josip Car, Imperial College London

Joanna Armstrong Schellenberg, London School of Hygiene and Tropical Medicine

Contact details

Barbara Willey, Barbara.Willey@lshtm.ac.uk

Acknowledgements

This systematic review is registered with the Institute of Education's Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), London, UK.

We are grateful for discussions with Dr Ruth Stewart, whose advice and experience on logic models was very helpful.

We are very appreciative to Dr Alexander Rowe for helpful discussions, and to him and his colleagues for sharing the results of their search strategy for their systematic review of interventions to improve health worker behaviour in LMICs, as presented at the Second Global Symposium on Health Systems Research, November 2012, Beijing, China.

We are also grateful to Dr Moriah Ellen and Dr Simon Lewin for helpful discussions and advice on how to prepare executive summaries for policy makers, as presented at the Second Global Symposium on Health Systems Research, November, 2012, Beijing, China.

Our sincere thanks to Dr Simon Lewin for sharing the SUPPORT collaboration executive summary template and allowing us to adapt this template for the executive summary of this systematic review.

Finally, many thanks to Dr Miguel A. González Block for his helpful peer review of this report.

Funding

This review was supported by the UK Department for International Development (DfID) through an open call for systematic reviews by AUSAID/3ie/DfID in November 2010.

Conflict of interest

We declare that we have no competing interests.

Appendix 1.2: Inclusion and exclusion criteria

Inclusion criteria

Population:

- Low and lower middle income economies as defined by the World Bank (see list in Appendix 2);
- Front-line health workers, i.e. health care personnel, including community health workers, delivering state-provided health care services to a targeted (e.g. child or maternal) or general population.

Interventions:

- Study described an intervention delivered at scale (in at least one district);
- Study included health services delivered by state providers in low or lower middle income countries;
- Study included a supply-side intervention to improve the health services provided by front-line workers during their interaction with users, for example, pre-service training, in-service training, supervision, guideline and protocol dissemination, reminders, quality improvement, quality assurance, audit and feedback, and checklists;
- Interventions which additionally addressed other constraints at the health service delivery level (i.e. shortage or distribution of staff; programme management and supervision; drug and medical supplies; equipment and infrastructure (Hanson et al., 2003), in addition to technical guidance (Table 1.1).

Outcomes:

- Study reported at least one of the following outcomes: survival impact, coverage, access, quality, equity.

Study design:

- Study was one of the following designs: RCT, CRT, CT, CBA or ITS.

Exclusion criteria

Interventions:

- Interventions targeting the introduction of new services, or the testing of novel delivery strategies;
- Interventions addressing supervision alone, where supervision was not one element within a multi-faceted intervention to improve the process of delivery of health services that met the inclusion criteria, as a comprehensive review of supervision in LMICs exists (Bosch-Capblanch et al., 2011).
- Interventions to strengthen health systems that were targeted at levels other than the service delivery level (i.e. community or household level; health sector policy and strategic management level; public policies cutting across sectors; or environmental and contextual characteristics), except where these also included interventions targeted at the health services delivery level (Hanson et al., 2003).

Appendix 2: Final MEDLINE search strategy

In order to arrive at this final search strategy, we used an iterative process within MEDLINE. We began with a very wide range of search terms, including key words and thesaurus terms. For example our first search (see Appendix 3: Initial MEDLINE search strategy) included thesaurus terms; AND keywords, including 'improve, increase, enhance'; AND developing countries; AND study designs. When limited to the review period (2000-2011) in MEDLINE, 5,297 records were returned. Screening 1,000 of these returned four records that met inclusion criteria, a success rate of 0.4 percent.

As double screening potentially over five thousand records from each of the three databases was not practical, and as we wished to retain the process of double screening due to the complex nature of the interventions, we took the pragmatic decision to refine our initial search strategy. Papers known to the review authors, as well as those identified during the initial screen, were used to refine the strategy by mapping how these known papers were indexed within MEDLINE. This allowed: some thesaurus terms to be removed; the removal of MESH term explosion from some thesaurus terms; the removal of free-text keywords (e.g. 'improve'); and the consolidation of terms relating to study design that also included the removal of free-text key words, e.g. 'comparison', as they overlapped in indexing with thesaurus terms and study design classifications that were retained.

The final MEDLINE search strategy is shown below.

1. "Delivery of Health Care"/or delivery of health care, integrated/

2. health personnel/ or allied health personnel/ or community health aides/ or nurses/ or pharmacists/ or physicians/

3. health services/ or community health services/ or child care/ or infant care/ or intensive care, neonatal/ or perinatal care/ or child health services/ or exp maternal health services/ or immunization programs/ or mass vaccination/ or vaccination/ or rural health services/

4. quality assurance, health care/

5. or/1-4

6. (randomized controlled trial or controlled clinical trial or clinical trial or evaluation studies or comparative study or multicenter study).pt.

7. research design/ or follow-up studies/ or prospective studies/ or cluster analysis/ or longitudinal studies/ or intervention studies/

8. 6 or 7

9. (letter or editorial or comment or review or case study or news).pt.

10. 8 not 9

11. (animals not (humans and animals)).sh.

12. 10 not 11

13. Angola/

14. Armenia/
15. Belize/
16. Bhutan/
17. Bolivia/
18. Cameroon/
19. Cape Verde/
20. China/
21. Congo/ or "Democratic Republic of the Congo"/
22. Cote d'Ivoire/
23. Djibouti/
24. Ecuador/
25. Egypt/
26. El Salvador/
27. "Georgia (Republic)"/ or Georgia/
28. Guatemala/
29. Guyana/
30. Honduras/
31. Indonesia/
32. India/
33. Iraq/
34. Jordan/
35. Micronesia/
36. Lesotho/
37. Indian Ocean Islands/
38. Moldova/
39. Mongolia/
40. Morocco/
41. Nicaragua/
42. Nigeria/
43. Pakistan/
44. Papua New Guinea/

45. Paraguay/

46. Philippines/

47. Samoa/

48. Atlantic Islands/

49. Senegal/

50. Sri Lanka/

51. Sudan/

52. Swaziland/

53. Syria/

54. Thailand/

55. East Timor/

56. Tonga/

57. Tunisia/

58. Turkmenistan/

59. Ukraine/

60. Uzbekistan/

61. Vanuatu/

62. Vietnam/

63. Yemen/

64. 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63

65. Afghanistan/

66. Bangladesh/

67. Benin/

68. Burkina Faso/

69. Burundi/

70. Cambodia/

71. Central African Republic/

72. Chad/

73. Comoros/

74. congo/ or "democratic republic of the congo"/

75. Eritrea/

76. Ethiopia/

77. Gambia/

78. Ghana/

79. Guinea/ or Equatorial Guinea/

80. Haiti/

81. Kenya/

82. "Democratic People's Republic of Korea"/

83. Kyrgyzstan/

84. Laos/

85. Liberia/

86. Madagascar/

87. Malawi/

88. Mali/

89. Mauritania/

90. Mozambique/

91. Myanmar/

92. Nepal/

93. Niger/

94. Rwanda/

95. Sierra Leone/

96. Melanesia/

97. Somalia/

98. Tajikistan/

99. Tanzania/

100. Togo/

101. Uganda/

102. Zambia/

103. Zimbabwe/

104. 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98 or 99 or 100 or 101 or 102 or 103

105. 64 or 104

106. developing countries/

107. 105 or 106

108. 5 and 12 and 107

109. limit 108 to yr="2000-2011"

Appendix 3: Initial MEDLINE search strategy

1. Angola/
2. Armenia/
3. Belize/
4. Bhutan/
5. Bolivia/
6. Cameroon/
7. Cape Verde/
8. China/
9. Congo/ or "Democratic Republic of the Congo"/
10. Cote d'Ivoire/
11. Djibouti/
12. Ecuador/
13. Egypt/
14. El Salvador/
15. "Georgia (Republic)"/ or Georgia/
16. Guatemala/
17. Guyana/
18. Honduras/
19. Indonesia/
20. India/
21. Iraq/
22. Jordan/
23. Micronesia/
24. Lesotho/
25. Indian Ocean Islands/
26. Moldova/
27. Mongolia/
28. Morocco/
29. Nicaragua/

30. Nigeria/

31. Pakistan/

32. Papua New Guinea/

33. Paraguay/

34. Philippines/

35. Samoa/

36. Atlantic Islands/

37. Senegal/

38. Sri Lanka/

39. Sudan/

40. Swaziland/

41. Syria/

42. Thailand/

43. East Timor/

44. Tonga/

45. Tunisia/

46. Turkmenistan/

47. Ukraine/

48. Uzbekistan/

49. Vanuatu/

50. Vietnam/

51. Yemen/

52. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51

53. Afghanistan/

54. Bangladesh/

55. Benin/

56. Burkina Faso/

57. Burundi/

58. Cambodia/

59. Central African Republic/

60. Chad/

61. Comoros/

62. congo/ or "democratic republic of the congo"/

63. Eritrea/

64. Ethiopia/

65. Gambia/

66. Ghana/

67. Guinea/ or Equatorial Guinea/

68. Haiti/

69. Kenya/

70. "Democratic People's Republic of Korea"/

71. Kyrgyzstan/

72. Laos/

73. Liberia/

74. Madagascar/

75. Malawi/

76. Mali/

77. Mauritania/

78. Mozambique/

79. Myanmar/

80. Nepal/

81. Niger/

82. Rwanda/

83. Sierra Leone/

84. Melanesia/

85. Somalia/

86. Tajikistan/

87. Tanzania/

88. Togo/

89. Uganda/

90. Zambia/

91. Zimbabwe/

92. 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91

93. 52 or 92

94. Developing Countries/

95. 93 or 94

96. health communication/ or health manpower/ or exp health personnel/ or health promotion/

97. "delivery of health care"/ or health/ or public health/ or health facilities/ or rehabilitation centers/ or allied health personnel/ or health services/ or economic development/ or "marketing of health services"/ or health planning/ or single-payer system/ or state medicine/ or regional health planning/ or health planning organizations/ or home care agencies/

98. (improv* or increas* or enhanc* or advanc* or ameliorat* or amend* or augment* or better or boost* or develop* or elevat* or help* or lift or polish or progress* or promot* or rais* or recover* or refin* or upgrad*).ab,ti.

99. (randomized controlled trial or controlled clinical trial or clinical trial or evaluation studies or comparative study).pt.

100. ("random*" or "placebo*" or trial or experiment* or intervention* or pre test or pretest or post test or posttest or preintervention or postintervention or time series or cross over or crossover or factorial* or latin square or assign* or allocat* or volunteer* or control* or compar* or prospectiv* or impact* or effect* or chang* or evaluat*).tw.

101. drug therapy.fs.

102. research design/ or follow-up studies/ or prospective studies/ or cross over studies/

103. groups.ab.

104. 99 or 100 or 101 or 102 or 103

105. 96 or 97

106. (animals not (humans and animals)).sh.

107. 104 not 106

108. 107 and 105 and 98 and 95

109. limit 108 to yr="2000-2011"

Appendix 4: Screening template

Type of study

Q1. Is this study any of:

- Individually-randomized controlled trials (RCT)
- Cluster randomized trials (CRT), including randomized stepped wedge designs
- Non-randomized cluster trials (CT) with at least two intervention sites and two control sites
- Controlled before-and-after studies (CBA), with at least two intervention sites and two control sites
- Interrupted time series studies (ITS) where the timing of the intervention is clear and at least three time points before and after the intervention are available
- YES: GO TO NEXT QUESTION NO: EXCLUDE UNCLEAR: SEE NEXT QUESTION

Participants

Q2. Is this study set in a relevant country?

- YES: GO TO NEXT QUESTION NO: EXCLUDE UNCLEAR: SEE NEXT QUESTION

Q3. Does this study include front-line health care workers from the public/government sector?

- YES: GO TO NEXT QUESTION NO: EXCLUDE UNCLEAR: SEE NEXT QUESTION

Interventions

Q4. Does the study include an intervention that aims to improve service delivery at the interface between front-line workers and health service users?

- YES: GO TO NEXT QUESTION NO: EXCLUDE UNCLEAR: SEE NEXT QUESTION

Q5. Does the study include an intervention for front-line health workers from state-provided health care services?

- YES: GO TO NEXT QUESTION NO: EXCLUDE UNCLEAR: SEE NEXT QUESTION

Q6. Does the study include an intervention that targets the supply-side of service provision, delivered at the operational level of health service delivery?

- YES: GO TO NEXT QUESTION NO: EXCLUDE UNCLEAR: SEE NEXT QUESTION

Q7. Is the intervention to improve delivery of health services implemented in at least one district (i.e. at scale)?

- YES: GO TO NEXT QUESTION NO: EXCLUDE UNCLEAR: SEE NEXT QUESTION

Outcomes

Q8. Does this study report at least one of the following outcomes:

- Impact (mortality or disease prevalence/incidence)
- Coverage
- Access
- Quality
- Equity

YES: INCLUDE NO: EXCLUDE UNCLEAR: REFER TO FULL PAPER FOR CLARIFICATION

Appendix 5: Data extracted

- Details of the publication;
- Study design, sample size, duration, timing of assessment;
- Country, scale of implementation;
- Characteristics of the study population, health workers and recipient population;
- Characteristics of the intervention and comparison groups, including: co-interventions, other health system strengthening interventions, health system level of implementation, front-line workers receiving intervention, implementers delivering the intervention, duration of intervention, length between intervention and evaluation, style and content of training, reported fidelity of implementation;
- Reported outcomes of survival impact, coverage, access and quality, including stratification of these by socio-economic status and other measures of equity where available;
- Risk of bias and quality of the evidence.

Appendix 6: Reasons for exclusion of papers at full screening

Reason for exclusion	Number of studies	References
Not an included intervention to strengthen delivery of services	10	Jokhio et al. (2005), Kumar et al. (2008), Matendo et al. (2011), Mbeba et al. (2011), Pagaiya and Garner (2005), Sharma et al. (2005), Skarbinski et al. (2009), Bhutta et al. (2011), Carlo et al. (2010), Wu et al. (2011)
Not at scale	3	Bhunia et al.(2010) , Darmstadt et al. (2010), Hermida and Robalino (2002)
Not included study design	3	Akter et al. (2009), Kelley et al. (2001), Naimoli et al. (2006)
Demand-side intervention	1	Lester et al. (2010)
Not health workers	1	Peters et al. (2006)
Not government facilities	1	Shah et al. (2007)
Not included outcome	1	Bryce et al. (2005)
Paper not available	1	Zhang et al. (2004)

Appendix 7: Suggested risk of bias criteria for EPOC reviews

Risk of bias for studies with a separate control group:

- Randomised controlled trials (RCTs);
- Non-randomised controlled trials (NRCTs);
- Controlled before-after (CBA) studies;

Nine standard criteria are suggested for all RCTs, NRCTs and CBA studies. Further information can be obtained from the Cochrane handbook section on risk of bias (Higgs and Green, 2011).

1. Was the allocation sequence adequately generated?

Score 'Low risk' if a random component in the sequence generation process is described (e.g., referring to a random number table). Score 'High risk' when a non-random method is used (e.g., performed by date of admission). NRCTs and CBA studies should be scored 'High risk'. Score 'Unclear risk' if not specified in the paper.

2. Was the allocation adequately concealed?

Score 'Low risk' if the unit of allocation was by institution, team or professional and allocation was performed on all units at the start of the study; or if the unit of allocation was by patient or episode of care and there was some form of centralised randomisation scheme, e.g., an on-site computer system or sealed opaque envelopes was used. CBA studies should be scored 'High risk'. Score 'Unclear risk' if not specified in the paper.

3. Were baseline outcome measurements similar? [5,6]

Score 'Low risk' if performance or patient outcomes were measured prior to the intervention, and no important differences were present across study groups. In RCTs, score 'Low risk' if imbalanced but appropriate adjusted analysis was performed (e.g. Analysis of covariance). Score 'High risk' if important differences were present and not adjusted for in analysis. If RCTs had no baseline measure of outcome, score 'Unclear risk'.

4. Were baseline characteristics similar?

Score 'Low risk' if baseline characteristics of the study and control providers were reported and similar. Score 'Unclear risk' if it is not clear in the paper (e.g. characteristics are mentioned in the text but no data were presented). Score 'High risk' if there is no report of characteristics in the text or tables or if there are differences between control and intervention providers. Note that in some cases

[5] If some primary outcomes were unbalanced at baseline, assessed blindly or affected by missing data and others were not, each primary outcome can be scored separately.

[6] If 'Unclear risk' or 'High risk', but there is sufficient data in the paper to do an adjusted analysis (e.g. Baseline adjustment analysis or Intention to treat analysis) the criteria should be re-scored as 'Low risk'.

imbalance in patient characteristics may be due to recruitment bias, whereby the provider was responsible for recruiting patients into the trial.

5. Were incomplete outcome data adequately addressed?[7]

Score 'Low risk' if missing outcome measures were unlikely to bias the results (e.g. the proportion of missing data was similar in the intervention and control groups or the proportion of missing data was less than the effect size, i.e. unlikely to overturn the study result). Score 'High risk' if missing outcome data were likely to bias the results. Score 'Unclear risk' if not specified in the paper (do not assume 100 percent follow-up unless stated explicitly).

6. Was knowledge of the allocated interventions adequately prevented during the study?[8]

Score 'Low risk' if the authors stated explicitly that the primary outcome variables were assessed blindly, or the outcomes were objective, e.g. length of hospital stay. Primary outcomes are those variables that correspond to the primary hypothesis or question as defined by the authors. Score 'High risk' if the outcomes were not assessed blindly. Score 'Unclear risk' if not specified in the paper.

7. Was the study adequately protected against contamination?

Score 'Low risk' if allocation was by community, institution or practice and it is unlikely that the control group received the intervention. Score 'High risk' if it is likely that the control group received the intervention (e.g. if patients rather than professionals were randomised). Score 'Unclear risk' if professionals were allocated within a clinic or practice and it is possible that communication between intervention and control professionals could have occurred (e.g. physicians within practices were allocated to intervention or control).

8. Was the study free from selective outcome reporting?

Score 'Low risk' if there is no evidence that outcomes were selectively reported (e.g. all relevant outcomes in the methods section were reported in the results section). Score 'High risk' if some important outcomes were subsequently omitted from the results. Score 'Unclear risk' if not specified in the paper.

9. Was the study free from other risks of bias?

Score 'Low risk' if there is no evidence of other risk of biases.

[7] If some primary outcomes were unbalanced at baseline, assessed blindly or affected by missing data and others were not, each primary outcome can be scored separately.

[8] If some primary outcomes were unbalanced at baseline, assessed blindly or affected by missing data and others were not, each primary outcome can be scored separately.

Appendix 8: Funding and affiliations of research study teams of included studies

Reference	Country and setting	Funder	Composition of research study team (based on author affiliations)
(A) Improving the service provided by front-line workers during their interaction with users			
Gilroy et al. (2004)	Mali; rural	United States Agency for International Development through the Family Health and Child Survival Cooperative Agreement (HRN-A-00-96-90006-00) with Johns Hopkins Bloomberg School of Public Health	Departments of Epidemiology and International Health, Johns Hopkins Bloomberg School of Public Health, USA Faculté de Médecine, de Pharmacie et d'Odontostomatologie, Université du Mali, Bamako, Mali Save the Children USA Save the Children, Bamako, Mali
Gulmezoglu et al. (2007)	Thailand and Mexico; peri-urban Thailand and urban Mexico	The trial was funded by the UNDP/UNFPA/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP)	UNDP/UNFPA/WHO/World Bank Special Programme on Research, Development and Research Training in Human Reproduction, Department of Reproductive Health and Research, World Health Organization, Switzerland Engenderhealth, New York, USA Khon Kaen University, Department of Obstetrics and Gynaecology, Faculty of Medicine, Khon Kaen, Thailand Ottawa Health Research Institute, Clinical Epidemiology Program, Canada
Ngasala et al. (2008)	Tanzania; rural	The study received financial support from the Swedish International Development Cooperation Agency SIDA/Sarec bilateral project between KI and MUCHS (Bil-Tz	Infectious Diseases Unit, Karolinska Institutet (KI), Sweden Department of Parasitology, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania Division of International Health, Karolinska Institutet, Sweden The Nordic school of Public Health, Gothenburg, Sweden

Reference	Country and setting	Funder	Composition of research study team (based on author affiliations)
		16/98 75007059)	Department of Pharmacology, Muhimbili University College of Health Sciences (MUCHS), Dar es Salaam, Tanzania Division of Clinical Pharmacology, Karolinska Institute, Sweden Global Malaria Programme, World Health Organization, Switzerland
Zaman et al. (2008)	Pakistan; urban	Department of Child and Adolescent Health of WHO that funded the Project with resources received for research from USAID and SAREC.	Health Services Academy, Islamabad, Pakistan Department of Social and Preventive Paediatrics, King Edward Medical College, Lahore, Pakistan Department of Child and Adolescent Health, World Health Organization, Switzerland
Zurovac et al. (2011)	Kenya; rural	The trial was funded by the Wellcome Trust, UK through a project grant to author Snow (grant number 084253), which also supported Zurovac. Snow is supported by the Wellcome Trust as Principal Research Fellow (grant number 079080)	Institute-Wellcome Trust Research Programme, Nairobi, Kenya Centre for Tropical Medicine, University of Oxford, UK Center for Global Health and Development, Boston University, USA Department of Disease Prevention and Control, Ministry of Public Health and Sanitation, Nairobi, Kenya Section of Infectious Diseases, Boston University School of Medicine, USA Malaria Branch, Division of Parasitic Diseases and Malaria, US Centers for Disease Control and Prevention, USA

Reference	Country and setting	Funder	Composition of research study team (based on author affiliations)
(B) Improving the service provided by front-line workers during their interaction with users, alongside strengthening at the service delivery level			
Ayieko et al. (2011)	Kenya; rural	Funds from a Wellcome Trust Senior Fellowship awarded to Mike English (#076827) supported intervention development, provision of guidelines, and job aids and all the research components. Routine hospital care was provided by the Government of Kenya.	KEMRI-Wellcome Trust Research Programme, Nairobi, Kenya Division of Child Health, Ministry of Public Health and Sanitation, Nairobi, Kenya Ministry of Medical Services, Nairobi, Kenya Department of Paediatrics and Child Health, University of Nairobi, Kenyatta National Hospital, Nairobi, Kenya Infectious Disease Epidemiology Unit, Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, United Kingdom Department of Paediatrics, University of Oxford and John Radcliffe Hospital, United Kingdom
Bradley and Igras (2005)	Guinea and Kenya; mixed urban and rural	UNICEF USAID	EngenderHealth, New York, USA
Jennings et al. (2010)	Benin; rural	This study was co-funded with resources received from the United States Agency for International Development (USAID) through the USAID Health Care Improvement Project (Contract No. GHN-I-0I-07-00003-00) and the Integrated Family Health	USAID Health Care Improvement Project, University Research Co., LLC, USA. Department of Population, Family, and Reproductive Health, Johns Hopkins Bloomberg School of Public Health, USA. Integrated Family Health Project, University Research Co., LLC, Bohicon, Benin.

Reference	Country and setting	Funder	Composition of research study team (based on author affiliations)
		Project (Contract No. 680-A-00-06-00013-00), both managed by University Research Co.	
Rowe et al. (2009)	Benin; mixed urban and rural	Funding was provided by the United States Agency for International Development's Africa Integrated Malaria Initiative (project 936-3100).	Malaria Branch, Division of Parasitic Diseases, Centers for Disease Control and Prevention, USA. Direction Départementale de la Santé Publique de l'Ouémé et Plateau [Departmental Direction of Public Health of Ouémé and Plateau], Ministry of Health, Porto-Novo, Benin. Africare-Benin, Porto-Novo. Division of General and Community Pediatric Research, Cincinnati Children's Hospital, USA. Data Management Activity, Division of Parasitic Diseases, Centers for Disease Control and Prevention, USA. Parasitic Diseases Branch, Division of Parasitic Diseases, Centers for Disease Control and Prevention, USA.
(C) Improving the service provided by front-line workers during their interaction with users, and strengthening at the service delivery, community/household or health sector policy level			
Arifeen et al. (2004a, b)	Bangladesh; rural	This study was done at the ICDDR,B: Centre for Health and Population Research, with funding from the Bill and Melinda Gates Foundation and through a grant to the WHO Department of	International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B), Dhaka, Bangladesh Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA Universidade Federal de Pelotas, Pelotas, Brazil

Reference	Country and setting	Funder	Composition of research study team (based on author affiliations)
		Child and Adolescent Health (Cooperative Agreement #388-A-00–97–00032–00) from the United States Agency for International Development.	
(Armstrong Schellenberg et al., 2004a, b)	Tanzania; rural	The Tanzania Essential Health Interventions Project is a Collaboration co-funded by the Tanzania Ministry of Health and the International Development Research Centre, Canada. The IMCI-MCE is arranged, coordinated and funded by the Department of Child and Adolescent Health and Development of the World Health Organization with the financial support of the Bill and Melinda Gates Foundation and the US Agency for International Development.	Ifakara Health Research and Development Centre, Ifakara, Tanzania London School of Hygiene and Tropical Medicine, UK Swiss Tropical Institute, Basel, Switzerland WHO Headquarters, Geneva, Switzerland International Development Research Centre, Ottawa, Canada Tanzania Essential Health Interventions Project, Ministry of Health, Dar es Salaam, Tanzania WHO Tanzania
Winch et al. (2003)	Mali; mixed urban and rural	This work was supported financially by the United States Agency for International Development	Johns Hopkins Bloomberg School of Public Health, USA Save the Children, Bougouni, Mali Faculté de Médecine, de

Reference	Country and setting	Funder	Composition of research study team (based on author affiliations)
		through the Family Health and Child Survival Cooperative Agreement (HRN-A-00-96-90006-00) with Johns Hopkins Bloomberg School of Public Health.	Pharmacie et d'OdontoStomatologie, Université du Mali, Bamako, Mali Secretary General, Ministry of Health, Bamako, Mali Save the Children USA, Westport, CT, USA

Appendix 9: Summaries of existing systematic reviews on interventions to improve health service delivery in LMICs

Shortage and distribution of appropriately qualified staff

Penaloza and colleagues (2011) found only one study from the Philippines in their review of interventions to reduce emigration of health workers from LMICs:

- This study reported increased emigration following changes in restrictions to immigration legislation in the USA.
- The authors concluded that 'no evidence was found regarding the effectiveness of interventions designed to decrease the emigration of health professionals, or the effectiveness of such interventions in LMICs in particular'.
- Specifically, no evidence was available on:
    - strategies to improve the working conditions or career prospects of health professionals;
    - compulsory service schemes;
    - strategies to facilitate and support the return of health professionals from abroad.
- The authors highlighted the need for robust studies to investigated the effectiveness of interventions to decrease the emigration of health professionals from LMICs.

Grobler and colleagues (2009) found no studies that met their study design inclusion criteria for interventions to recruit and retain health workers in rural areas of HICs or LMICs:

- The authors made some observations based on low-quality evidence:
    - Observational studies from high-income countries suggested that health professionals from rural backgrounds were more likely to practise in rural areas;
    - Evidence from four quasi-randomised trials in high-income countries suggested that clinical rotations a rural setting might increase medical students' decisions to practise in a rural area;
    - Descriptive studies of compulsory placement in Ecuador and South Africa were inconclusive;
    - No evidence on financial or non-financial incentives to health professionals was identified from LMICs, and evidence from high-income settings was mixed.
- The authors highlighted the need for robust studies to investigate the effectiveness, applicability and cost-effectiveness of strategies to recruit and retain health workers in rural areas of LMICs.

Sibley and colleagues (2007) included four studies assessing the effects of additional training for traditional birth attendants in LMICs:

- The authors concluded that:
    - it was unclear whether there was an impact on maternal mortality from moderate quality evidence from one study in Pakistan;
    - there was mixed evidence of low quality on maternal morbidity from Pakistan and Malawi;
    - there was moderate-quality evidence of about 30 percent reductions in perinatal mortality, neonatal mortality and stillbirths from three studies in Pakistan;
    - there were mixed findings of moderate quality on infant feeding advice from one study in Bangladesh;
    - There were mixed findings of very low quality for referral of mothers with complications from one study in Pakistan.
- The authors highlighted that no evidence on cost or cost-effectiveness was available and that numerous context issues needed to be considered when determining whether intervention effects are likely to be transferrable to other settings.

Lewin and colleagues (2010) included 82 studies, 16 from LMICs assessing the effect of lay health worker interventions in improving maternal and child health outcomes:

- The authors concluded that:
    - there was evidence of moderate quality to suggest improved uptake of immunisation among children <2 years of age, but only from high-income countries;
    - there was evidence of moderate quality, including from LMICs in Asia and Latin America, to suggest improved breastfeeding practices;
    - there was low-quality evidence from LMICs to suggest a reduction in under-five mortality of about 25 percent;
    - there was low-quality evidence from LMICs to suggest a reduction in neonatal mortality of about 24 percent;
    - there was low-quality evidence from LMICs to suggest a reduction in child morbidity of about 14 percent;
    - there was low-quality evidence from LMICs to suggest an improvement in care seeking practices for sick children of 30 percent;
    - there was moderate-quality evidence to suggest improved TB cure rates although little or no difference to completion of TB treatment, including from LMICs in Africa and Asia.
- The authors highlighted that little evidence was available in relation to substituting lay health workers for health professionals, and that there were numerous context issues to be considered when determining the transferability of findings to other settings.

Weak technical guidance

Opiyo and English (2010) included two studies, one each from Kenya and Sri Lanka in their review assessing the effectiveness of in-service training for seriously ill newborns and children:

- The authors concluded that:
    - there was moderate-quality evidence for an improvement in resuscitation behaviour from the Kenyan study;
    - There was moderate-quality evidence that there was little or no difference in neonatal mortality from the Kenyan study.
- The authors recommended that the impact of in-service training on long-term outcomes should be evaluated, and that training should be evaluated in terms of the costs and resources required.

Forsetlund and colleagues (2009) included 81 studies, with only 12 from LMICs, investigating the effect of educational meetings (lectures, workshops, or courses) that were sometimes part of a package of interventions with reminders, feedback reports and educational outreach:

- The authors concluded that:
    - there was moderate-quality evidence to suggest that educational meetings alone or combined with other interventions could improve professional practice; however, the effect was most likely to be small;
    - educational meetings alone were not likely to be effective for changing complex behaviour.

Van Lonkhuijzen and colleagues (2010) included 38 studies from LMICs in their review of training programmes aimed at improving emergency obstetric care among skilled birth:

- The strength of evidence was not categorised using the GRADE criteria; however, fewer than a quarter of the studies were robust designs, and therefore the quality of the evidence is likely to be low.
- The authors concluded that the majority of the studies reported improved knowledge and skills, but that this for the most part did not translate to improved behaviour:
    - It was not possible to comment on the role of hands-on clinical practice versus simulated training;
    - There was some suggestion that team approaches were beneficial;
    - The authors observed that the healthcare settings in which professionals worked was important.
- The authors observed that because this review did not restrict by study design, evaluation was challenging, due to inadequate study design and the use of non-validated measuring instruments.

Weak programme management

Jamtvedt and colleagues (2003) assessed the effects of audit and feedback on the practice of healthcare professionals. Their review was restricted to RCTs. They included 118 studies, only four of which were from LMICs:

- The authors concluded that:
    - there was low quality evidence, principally from high-income countries, to suggest increased compliance with desired behaviour and that the effects were generally small to moderate;
    - the effects were larger when baseline compliance to recommended practice was low and feedback was provided more intensively;
    - the evidence did not support the mandatory use of audit and feedback as an intervention to change provider practice.
- The authors highlighted that numerous context issues needed to be considered when determining whether intervention effects were likely to be transferrable to LMICs, including what current compliance with guidelines was like, the feasibility of using routinely collected data, and whether small improvements would be worthwhile.

Weak supervision

Bosch-Capblanch and colleagues (2011) reviewed the evidence of the effect of managerial supervision of health workers on the quality of primary health care in LMICs. Supervision was defined as site visits from a central level of the health system, plus at least one supervisory activity. This review was restricted to robust designs (RCT, CBA, ITS), and nine studies from LMICs in Africa, Asia and Latin America were included:

- Most outcomes were scores relating to providers' practice and knowledge, and provider or user satisfaction.
- The authors concluded that it was not known whether substantive, positive effects on the quality of primary healthcare in LMICs were achieved, with studies showing mixed results. Some low-quality evidence suggested that:
    - in comparison to no supervision, small benefits on provider practice were found, based on two studies with low-quality evidence;
    - in comparison to no supervision, one study showed small benefits on provider knowledge while another was inconclusive; both had low-quality evidence;
    - in comparison to no supervision, one study showed small benefits on drug stock management; this study had very low-quality evidence;
    - among studies comparing methods of providing supervision, frequent supportive supervision demonstrated small benefits on workers' performance from two studies with very low-quality evidence;

    - there was no evidence that reducing the frequency of visits from monthly to quarterly visits (over a nine-month study period) affected utilisation of services, from one study that compared the impact of less-intensive supervision, but again, the quality of the evidence was low.

- The authors highlighted that there was insufficient high-quality evidence to advocate for any particular form of implementing supervision, although there was no suggestion that more frequent supervision (monthly) was more beneficial than quarterly visits combined with supplementary visits to deal with emergencies or improve staff performance.

- The authors also suggested that future evaluations comparing supervision strategies should be 'guided by a framework that would allow identification of the key components of supervision that are linked with positive outcomes and the resources needed to implement these' (p.20).

Weak technical guidance together with programme management and supervision

Smith and colleagues (2009) included 16 studies from 13 malaria-endemic African countries, investigating provider-side interventions to improve prompt and effective treatment of fever in children under five years of age. This review included RCT, ITS, CBA and uncontrolled before-and-after studies, as well as controlled after studies. Public and private-sector and retail providers of antimalarials were included:

- The authors did not provide a summary of the quality of the evidence; however, it is likely that the available quality of the evidence was very low or low, given the designs of the included studies.

- Only two studies provided information on full Roll Back Malaria indicators of prompt and effective treatment, and although these showed improved practices, only about 30 percent coverage was achieved.

- The authors highlighted that most of the included studies evaluated the knowledge of providers rather than their practices. Some findings from probably low- or very low-quality evidence suggests that:

    - in-service training did not appear to have had much impact;

    - less didactic training e.g. IMCI, showed better results;

    - a combination of in-service training with job aids and frequent supervision showed better results;

    - there was some evidence that pre-packaging antimalarials might improve adherence to guidelines.

- The authors highlighted that evidence on the sustainability of any changes in knowledge or practice was lacking.

- The authors highlighted that assessment of effectiveness was limited by study designs.

Oliveira-Cruz and colleagues (2003) reviewed the evidence for interventions aiming to overcome constraints related to the delivery of health services at the close-to-client level in LMICs. Eligible interventions included management system

strengthening, training, supervision and the introduction of quality assurance systems. No restrictions on study design were applied to this review; however, opinion pieces were excluded. As such, this review included studies that only described an intervention as well as studies that reported the evaluation of an eligible intervention. In total 56 studies from LMICs were:

- The authors did not report the effect sizes of evaluations or detail what subset of their included studies were descriptive and what subset were intervention studies.
- It is not possible to determine the quality of evidence available upon which the conclusions were based. The authors reported that:
    - interventions to improve health worker skills (training, supervision, quality assurance) showed mixed effects in improving health worker knowledge;
    - training interventions targeting prescribing practices were generally positive
    - quality assurance interventions improved waiting times on the whole;
    - management strengthening interventions did not tend to report outcomes of service utilisation or coverage; however, process indicators such as rational use of resources, data collection and use, and staff morale were suggested to have improved.
- The authors highlighted that the studies included rarely allowed assessment of effectiveness due to limitations in study design and implementation.
- The authors highlighted that few studies presented the costs of the intervention.
- The authors highlighted that there was a need for future research to identify hindering and facilitating factors within the context, in addition to effectiveness.
- The authors highlighted that the evidence was insufficient with respect to equity and whether interventions were successfully reaching the poor and disadvantaged.

Rowe and colleagues (2012), in an ongoing review have presented preliminary results of a systematic review of the effectiveness and costs of strategies to improve health worker performance in low- and middle-income countries. This review includes interventions targeting constraints at the community and household level, at the health service delivery level, and at the health sector policy and strategic management level, and as such is extremely inclusive. RCTs, ITS, CBA and after-only studies with a randomised control group were included. The authors have over the past six years searched 15 electronic databases and 29 document inventories, screened >105,000 citations and included ~500 studies, about 25 percent of which are randomised controlled trials.

- The authors have not summarised the quality of the evidence available.

- The authors have summarised effects of the absolute changes in process outcomes of health worker practice using a median effect size.
- Many (40 percent of studies) show a small effect size of <10 percentage point increases.
- Interim results suggest that interventions with numerous components (e.g. training + job aids + supervision + community activities) have improved effects and larger median effect sizes in comparison to studies with only one component.
- Studies that also included a community activity component also seem to have improved effects.
- Future analysis will incorporate costs and investigate factors associated with increased effectiveness.

The authors of this report were supported by the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre).

The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) is part of the Social Science Research Unit (SSRU), Institute of Education, University of London.

Since 1993, we have been at the forefront of carrying out systematic reviews and developing review methods in social science and public policy. We are dedicated to making reliable research findings accessible to the people who need them, whether they are making policy, practice or personal decisions. We engage health and education policy makers, practitioners and service users in discussions about how researchers can make their work more relevant and how to use research findings.

Founded in 1990, the Social Science Research Unit (SSRU) is based at the Institute of Education, University of London. Our mission is to engage in and otherwise promote rigorous, ethical and participative social research as well as to support evidence-informed public policy and practice across a range of domains including education, health and welfare, guided by a concern for human rights, social justice and the development of human potential.

This material has been funded by the Department for International Development. However the views expressed do not necessarily reflect the department's official policies.

The report was first published in 2013 by:

Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)
Social Science Research Unit
Institute of Education, University of London
18 Woburn Square
London WC1H 0NR
Tel: +44 (0)20 7612 6397 http://eppi.ioe.ac.uk http://www.ioe.ac.uk/ssru

ISBN: 978-1-907345-55-5

This document is available in a range of accessible formats including large print.
Please contact the Institute of Education for assistance:
telephone: +44 (0)20 7947 9556 email: info@ioe.ac.uk