BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

A Systematic Review of Interventions for Improving the Performance of Community Health Workers in Low- and Middle-Income Countries

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-014216
Article Type:	Research
Date Submitted by the Author:	08-Sep-2016
Complete List of Authors:	Ballard, Madeleine; University of Oxford, Department of Social Policy and Intervention Montgomery, Paul; University of Oxford, Centre for Evidence Based Intervention
Primary Subject Heading:	Global health
Secondary Subject Heading:	Evidence based practice, Health policy, Public health
Keywords:	Human resource management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, PUBLIC HEALTH, Community child health < PAEDIATRICS, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts

A Systematic Review of Interventions for Improving the Performance of Community Health Workers in Low- and Middle-Income Countries

Madeleine Ballard, BA, MSc[a]
Paul Montgomery, BA, MSc, DipSw, DPhil[a]

[a] Centre for Evidence-Based Intervention, University of Oxford
32 Wellington Square, Oxford OX1 2ER, UK

For peer review only

Correspondence:

Madeleine Ballard
Centre for Evidence-Based Intervention
University of Oxford
32 Wellington Square
Oxford OX1 2ER, UK
E: madeleine.ballard@gmail.com
P: +447454266672

ABSTRACT

Objective: To systematically review and critically appraise the evidence for the effects of interventions to improve the performance of community health workers (CHWs) for community-based primary healthcare in low- and middle-income countries.

Design: Systematic review following PRISMA guidelines.

Methods: Nineteen electronic databases (e.g., CENTRAL, MEDLINE) were searched with a highly sensitive pre-specified strategy supplemented by cross-referencing, handsearching, and contacting experts completed July 2016. Randomised controlled trials evaluating interventions to improve CHW performance in LMICs were included and appraised for risk of bias. Outcomes were biological and behavioural patient outcomes (primary), utilisation of health services, quality of care provided by CHWs, and CHW retention (secondary).

Results: Two reviewers screened 8082 records; fourteen evaluations were included. Due to heterogeneity and lack of clear outcome data, no meta-analysis was conducted. Results were presented in a narrative summary. The review found one high-quality study showing no effect on the biological outcomes of interest, though these may not be indicative of all biological outcomes. It also found moderate quality evidence of the efficacy of performance improvement interventions for (i) improving behavioural outcomes for patients, (ii) improving utilisation of services by increasing the absolute number of patients who access services and, perhaps, better identifying those who would benefit from such services, and (iii) improving CHW quality of care in terms of upstream measures like completion of prescribed activities and downstream measures like adherence to treatment protocols. Nearly half of studies were compound interventions, making it difficult to isolate the effects of individual performance improvement intervention components, though four specific strategies pertaining to recruitment, supervision, incentivisation, and equipment were identified.

Conclusions: Variations in recruitment, supervision, incentivisation, and equipment may improve CHW performance. Practitioners should, however, assess the relevance and feasibility of these strategies in their health setting prior to implementation.

Keywords: Community health workers; lay health workers; supervision; incentives; recruitment; training; systematic review

Strengths and limitations of this study

- This is the first systematic review of interventions for improving the performance of community health workers that conducts risk of bias assessment, assessment of the quality of the body of evidence, and follows the PRISMA reporting guidelines.
- The review strengths include its pre-specified, systematic, and highly sensitive search strategy; inclusion of published and unpublished literature; and inclusion of only the most rigorous evidence.

For peer review only

- Though the review has carefully specified the inclusion criteria in keeping with previous studies, there is no widely accepted definition for this cadre of health worker and so some exclusions may be debated.
- Poor reporting and nonresponse from contacted authors meant it was not possible to obtain some relevant missing data on methodological characteristics, clinical characteristics, and outcomes for some trials
- Given mounting pressure to meet ambitious new international health goals and avoid repeating the mistakes of large-scale CHW programs of the past, the review is timely.

INTRODUCTION

Community health workers (CHWs)—lay workers to whom simple medical procedures can be "task shifted" from doctors—were widely promoted as a means to provide primary healthcare in resource poor settings as early as the 1978 Alma-Ata Declaration [1]. A series of reviews in the late 1980s and early 1990s, however, found that large-scale CHW programs often failed to replicate the success of smaller community-based programs [2-7].

Prompted by the failure to meet the health-related Millennium Development Goals [8, 9] and a continuing global health worker shortage that imperils the new Sustainable Development Goals (SDGs) [10], CHWs have once again been proposed as a way to extend services to hard-to-reach populations in remote areas [11-15].

During this period, rigorous evidence accumulated on the efficacy of CHWs to deliver assorted health interventions [16-19]. Yet the most recent evaluations of national-scale CHW programs remain unfavourable [20-22].

Statistical significance and optimisation are distinct and independent concepts and methods in intervention science have largely been developed to measure the former [23]. Without more insight into intervention components, it is difficult to assess whether a given large-scale CHW program did not achieve a statistically significant effect because (i) it does not work, (ii) it was not implemented properly

For peer review only

(type III error), or (iii) it was not yet optimised [23-26]. Given that qualitative assessments suggest it is often the latter [2, 27-32] disaggregating the impact of individual intervention components is warranted [23].

While history has demonstrated that broad contextual factors related to the health system or the political and economic climate can have an impact on community health worker performance [33], this review is focused on how modifiable aspects of the intervention itself—CHW characteristics (e.g. literacy) and program features (e.g. recruitment, supervision)—can be used to optimise CHW performance.

METHODS

Inclusion criteria

Because health workforce outcomes are empirically testable and workplace interventions possible to randomise, eligible trials were randomised controlled trials, cluster-randomised trials, crossover trials, and factorial trials. Studies were only included where the primary objective, or one of the primary objectives, was to determine the efficacy or effectiveness of an intervention to improve the performance of CHWs. This includes studies like 'head-to-head' comparisons of CHWs delivering the same intervention with different types of support that have to date been excluded from reviews of CHW effectiveness [e.g. 18].

Several definitions for and variants of the term community health worker have been employed in the literature. For the purpose of this review, CHWs were defined as: any lay health workers who (a) live in the catchment they serve, (b) are primarily based in the community (as opposed to a health facility), (c) belong to the formal health system (i.e. are managed by the government or an implementing NGO), (d) perform interventions related to healthcare delivery, and (e) have received organised training but have no formal or paraprofessional certification or tertiary education

For peer review only

degree [18, 34]. Aside from residence in a low- or middle-income country (LMIC) as classified by The World Bank Group (listed in Appendix B), there were no restrictions on the types of patients for whom data were extracted.

Any intervention designed to improve CHW performance, compared to CHWs who did not experience the intervention, was included. Studies that were designed to examine the efficacy of a particular CHW activity (e.g. zinc tablets for diarrhoea), as opposed to the relative effectiveness of different ways of supporting CHWs to deliver said activity were excluded. Interventions of any duration or follow up were included.

A number of processes and outcomes might be affected by interventions that aim to improve the performance of CHWs. Outcomes were extracted and categorised as biological and behavioural patient outcomes (primary), utilisation of health services, quality of care provided by CHWs, and CHW retention (secondary). Outcomes of interest were not used as criterion for screening studies.

Search Strategy

A comprehensive search strategy was used to identify all relevant studies regardless of language, year, or publication status (published, unpublished, in press, and in progress). We searched:

1. 3ie Impact Evaluation Database
2. AMED
3. BiblioMap
4. British Library for Development Studies at IDS
5. CINAHL
6. CENTRAL
7. DoPHER
8. EMBASE
9. Global Health
10. IDEAS
11. IPA Database
12. Databases within ISI Web of Science
13. JOLIS
14. J-PAL Evaluations Database
15. MEDLINE
16. MEDLINE In-Process & Other Non-Indexed Citations
17. POPLINE
18. Pubmed
19. World Bank IEG
20. WHOLIS

For peer review only

The following strategy was refined in consultation with an information specialist and used to search MEDLINE. In order to maximise sensitivity, no randomisation filter was included [35]. (Precise dates of coverage, date last searched, search strategies used for the databases listed above, and details of other resources searched can be found in Appendix C.)

1. ((community or village) adj2 (agent$ or aid$ or promot$ or mobili?er$ or distribut?r$ or worker$)).ti,ot,ab,kf.
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliary$ or extension) adj2 (worker$ or volunteer$)).ti,ot,ab,kf.
3. (accompanier$ or accompagnateur$ or activista$ OR animatrice$ OR brigadista$ or kader$ or monitora$ or promotora$ or sevika$ or fhw$ or chw$ or lhw$ or vhw$ OR chv$ or "shastho shebika" or "shasto karmis" or anganwadi$ or "barefoot doctor" OR "agente comunitario de salud" or "agente communitario de saude").ti,ot,ab,kf.
4. (performance or effective$ or skill$).ti,ot,ab,kf.
5. 1 or 2 or 3
6. 5 and 6

Searches were conducted May 4, 2014 and updated ending July 12, 2016.

Data collection and analysis

Two authors independently screened titles, abstracts, and full text articles. Included articles were determined by consensus. The extraction sheet was informed by the Cochrane EPOC Group data collection checklist [36], the Template for Intervention Description and Replication [37], and the Oxford Implementation Index [38] (see Appendix D). Reviewers extracted data on study design, participants, interventions, outcomes and methodological quality.

Risk of bias for all RCTs was assessed using The Cochrane Collaboration's tool for evaluating the risk of bias [39]. The quality of evidence across a body of evidence (i.e., multiple studies with similar interventions and outcomes) was assessed using the GRADE approach [40]. These assessments are made with the caveat that—as had been noted elsewhere [41-43]—there are issues with the applicability,

For peer review only

reproducibility, and clarity of GRADE in the field of public health and with complex interventions.

Cluster-randomised trials were verified to ensure appropriate analysis had been done (i.e. adjustment for clustering); if such an analysis had not been done, the necessary intracluster correlation coefficients (ICCs) were extracted or obtained from the investigators and results reanalysed [39]. All comparisons from factorial studies were included.

RESULTS

Characteristics of included studies

Excluding duplicates, a total of 8082 records were screened for inclusion (Fig 1). A total of 12 records reporting 14 studies were included [44-55]. Complete information for each trial is provided Appendix E; study characteristics are described in Table 1 [37]. Information on excluded, ongoing and studies awaiting classification is provided in Appendix F.

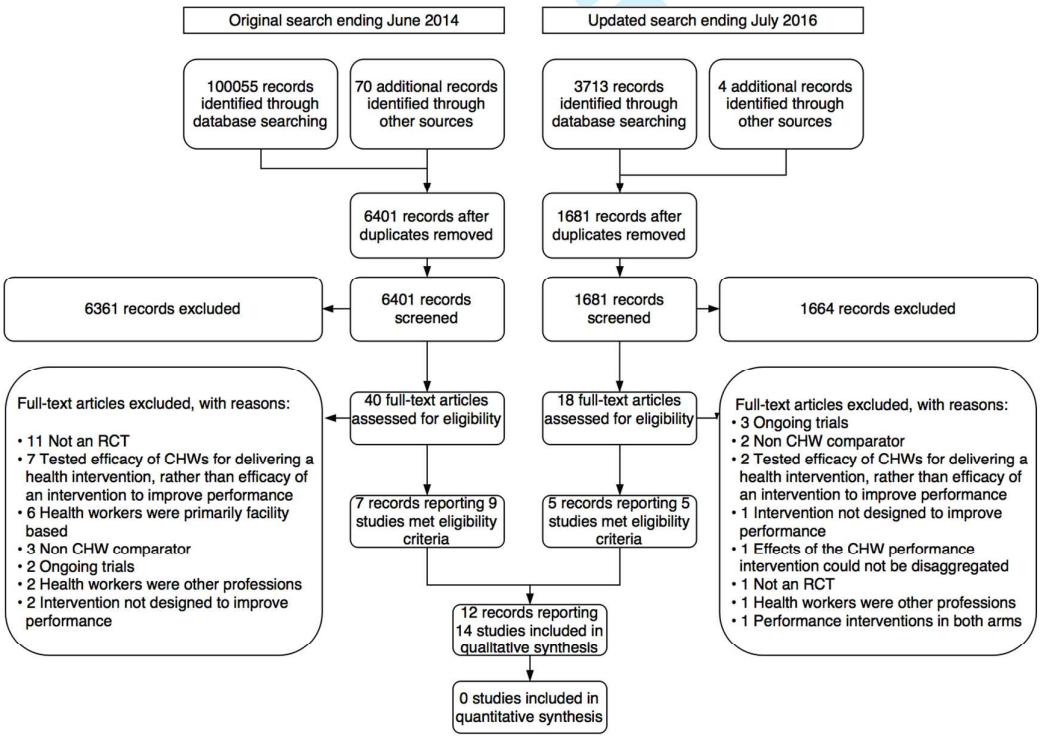


Figure 1 PRISMA study flow diagram

Ten countries were represented: Bangladesh, Ethiopia, Guatemala, India (2), Mali, Pakistan (2), Paraguay, Uganda, Tanzania (3), Zambia. Five trials evaluated compound interventions [49-53], three supervision interventions [45ABC], two incentive interventions [48, 55], two equipment interventions [44, 46], and one a recruitment intervention [54] and a training intervention [47]. No information about implementation fidelity was provided. The reporting of CHW and patient characteristics and contextual factors was poor. Assessments of the risk of bias for included studies are included in Appendix G and are considered when assessing the findings.

Table 1 Brief Study Characteristics

Study	Setting/ Country (Study Design)	CHW & Patient Sample	Intervention Type & Description of Trial Arms	Outcomes (Follow-up)
Andre-oni 2016 [48]	**Objective:** Test the effect of tailored contracts for CHWs conducting door-to-door immunisation drives in rural and urban Pakistan			
	Rural & Urban Pakistan (2x2 factorial RCT)	Female CHWs (I_{a+c}=85, I_{a+d}=84, C_{b+c}=88, C_{b+d}=80). Patients were children under 5 yrs. (N= unspecified)	Incentives I: a+c; a+d C: b+c; b+d (a) Employment contracts tailored to individual discounting preferences (b) Untailored contracts (c) Advanced allocation of daily work targets (d) Day-of allocation	CHWs: Closeness to vaccine distribution policy objective Patients: None (Immediate)
Ashraf 2015 [54]	**Objective:** Test the effect of advertising career possibilities at the recruitment stage for CHWs providing basic primary care			
	Rural Zambia (C-RCT)	CHWs (I=149, C=149). Patients were children under 5 yrs. & women (N = unspecified)	Recruitment I: Recruitment materials emphasising career possibilities C: Recruitment materials emphasising benefits to community	CHWs: Home visits, patients seen, comm. mtgs organised Patients: Use of services (18 months)

Ayele 1993 [53]	**Objective:** Test the effect of an in-person refresher course and monthly "supportive" supervision on CHWs conducting health promotion			
	Rural Ethiopia (C-RCT)	Male CHWs (I=50, C=50). Patient characteristics not reported (N= unspecified)	Training-Supervision I: 5 day in-person refresher course and monthly supervision C: No refresher, no supervision	CHWs: Composite functional status score Patients: None (6 months)
Bailey 1996 [47]	**Objective:** Test the effect of tiered, interactive training featuring visual cue cards in improving CHW case management of diarrhoea			
	Rural Guatemala (RCT)	CHWs (I=22, C=27). Patient characteristics not reported (N= unspecified)	Training I: Interactive, tiered training with visual cue cards C: Didactic training	CHWs: Knowledge score Patients: None (Immediate)
Bossu-roy 2016 [55]	**Objective:** Test the effect of performance-based incentives on CHWs providing TB case detection and adherence counselling			
	Urban India (C-RCT, 4 arm)	CHWs (N=78). Adult patients (N = 2500).	Incentives I_a: Financial incentives based on patient detection for six months then based on treatment adherence I_b: Financial incentives based on patient detection for six months then a fixed salary I_c: Fixed salary for six months then incentives based on treatment adherence C: Fixed salary	CHWs: TB case detection Patients: TB default rate (6 months)
Carrón 1994 [52]	**Objective:** Test the effect of an interactive counselling guide and supervisory behavioural feedback instrument on CHWs distributing family planning materials			
	Rural Paraguay (2x2 factorial C-RCT)	Female CHWs (I_a=45, I_b=45, I_{a+b}=45, C=45). Patient characteristics not reported (N=unspecified)	Training-Supervision I: a; b; a+b (a) Trained using interactive service algorithm (Decision tree) (b) Role-play based supervision (Behavioural feedback instrument) C: Usual training and supervision	CHWs: Quality of care provided Patients: None (12 months)
Chang 2011	**Objective:** Test the effect of an mobile phone CHW support intervention on HIV care			

[50]	Rural Uganda (C-RCT)	Female CHWs (I=13, C=16). Adult HIV+ patients (I=446, C=524)	Integration-Equipment I: A text message reporting system and health centre hotline C: No mobile health intervention (usual care)	CHWs: None Patients: Virologic failure, adherence (2 years)
DeRenzi 2012A [45]	**Objective:** Test the effect of SMS reminders to improve the promptness of routine CHW visits			
	Urban Tanzania (Pilot RCT)	Female CHWs (I=8, C=7). Patient characteristics not reported (N= unspecified)	Supervision I: Escalating text message reminder system and follow-up for underperforming CHWs C: Usual care	CHWs: None Patients: Change in closed referrals (24 days)
DeRenzi 2012B [45]	**Objective:** Evaluate the impact of SMS reminders plus escalation to supervisor to improve the promptness of routine CHW visits			
	Urban Tanzania (RCT)	CHWs (I=34, C=40) Chronic care patients (N=unspecified)	Supervision I: Escalating text message reminder system & follow-up for underperforming CHWs C: Usual care	CHWs: None Patients: Mean time overdue at clinic (40 days)
DeRenzi 2012C [45]	**Objective:** Test the effect of escalating to the supervisor versus SMS messages alone in the context of SMS reminders to improve the promptness of routine CHW visits			
	Urban Tanzania (RCT)	CHWs (N=87) Chronic care patients (N=unspecified)	Supervision I: Escalating text message reminder system only C: Usual care	CHWs: None Patients: Mean time overdue at clinic (3 months)
Gauth-am 2015 [44]	**Objective:** Test the effect of mobile phone-based, media-rich procedural guidance applications on CHWs providing basic primary care			
	Rural India (RCT)	CHWs (I=8, C=8). Paediatric and adult patients (I=65, C=61)	Equipment I: Mobile phone–based, media-rich procedural guidance applications C: Paper guidelines	CHWs: Protocol compliance Patients: None (2 months)
Gins-burg 2014	**Objective:** Test the effect of mobile guidance applications on CHWs referring women with abnormal breast exams			
	Rural	Female CHWs	Equipment (I_a)/	

[46]	Bangla-desh (3-arm RCT)	(I_a=10, I_b=10, C=10). Female adult patients (I_a=246, I_b=240, C=70)	Training-Equipment (I_b) I: a; b (a) Smart phone with applications to guide interview, report data, show motivational video, and offer an appointment (b) As in (a) plus additional "patient navigation training." C: Paper-based tools	Patient adherence with referral (1 month)
Omer 2008 [49]	**Objective:** Test the effect of an indigenous, evidence-based communication tool on CHWs providing maternal care			
	Urban Pakistan (C-RCT)	Female CHWs (I=52, C=unspecified). Mothers and their children (I=529, C=541)	Integration-Training I: Indigenous, evidence-based communication tool C: Usual care	CHWs: Home visits Patients: Uptake of protective health practices (10 months)
Winch 2003 [51]	**Objective:** Test the effect of visual aids, refresher training, a counter referral recordkeeping notebook, and a community meeting on CHWs treating common illnesses			
	Rural and Urban Mali (C-RCT)	CHWs (N=~102). Children <14 yrs. (N=152, C=134)	Integration -Training-Reporting- I: Visual aids, refresher training, a counter referral recordkeeping notebook, a community meeting C: Standard training	CHWs: Proportion of children treated at home Patients: Closed referrals

$\bar{x}$**:** mean; **C**: control; **Comm:** community; **C-RCT:** cluster RCT; **I:** intervention; **N:** intervention/control participant breakdown unspecified, total number of participants; **mtg:** meeting; **RCT:** randomised controlled trial; **yrs**: years

Effects of Interventions

The protocol for this review (see Appendix A) specified that a meta-analysis would be conducted if appropriate. Meta-analysis would, however, be uninformative due to heterogeneity and lack of clear outcome data. Results are presented in a structured summary [38]. Where possible, standardised measures of effect (RR, SMD) have been provided to aid comparison. Due to missing data, multiple outcome

For peer review only

measures were not possible to standardise and are reported as in the original trial papers. Although six authors were contacted regarding missing data, none were able to supply it.

Five studies examined the review's primary outcomes [45BC, 49-51], (Table 2) and nine studies examined secondary outcomes [44, 45A, 46-49, 52-54] (Table 3).

Primary outcomes

The primary outcomes specified in the protocol were objective physical outcomes for patients and subjective health outcomes for patients.

Only one study reported biological outcomes for patients [50], finding no significant differences in HIV patients' cumulative risk of virologic failure (RR = 1.17, 95% CI 0.84–1.64, P = 0.34) or mortality (RR = 0.82, 95% CI 0.55–1.22, P = 0.33). This evidence of no effect was graded high (no reason for downgrade).

Table 2 Primary outcomes

Primary outcome measure	Study (N)	Intervention Type	Variables	Measure of Effect (95% CI or SE)
Biological outcomes for patients	Chang 2011 (N=970 patients)	Reporting-Supervision	Virologic failure	RR = 1.17 (0.84–1.64) P = 0.34
			Mortality	RR = 0.82 (0.55–1.22) P = 0.33
Behaviou-ral outcomes for patients	Bossuroy 2016 (N = 2500 patients)	Incentives	TB default rate	B = -0.04 (SE = 0.06) P > 0.05
	Chang 2011 (N=970 patients)	Reporting-Supervision	<95% pill adherence	RR = 1.01 (0.97–1.06) P = 0.59
			Loss to follow-up	RR = 1.29 (0.50–13.32) P = 0.60
	DeRenzi 2012B	Supervision	Mean days clients overdue	U = 271.00 P < 0.001

(N=un-specified patients)			r = 0.500
DeRenzi 2012C (N=un-specified patients)	Supervision	Mean days clients overdue	See text
Omer 2008 (N=1070 patients)	Integration-Training	Attend at least one prenatal visit	RR = 1.94 (1.56-2.41) P < 0.0001
		Colostrum to newborns	RR = 1.21 (1.12-1.32) P < 0.0001
		Stop routine heavy work	See text
		Exclusive breastfeeding	See text
Winch 2003 (N=286 patients)	Training-Reporting-Integration	Proportion of malarial children treated by mothers exactly per IMCI norms	RR = 27.76 (0.53-1441.23) P < 0.0001

Six studies reported behavioural outcomes for patients: Bossuroy [55] found no significant difference in patient TB default rates (B = -0.04, SE = 0.06, P > 0.05), though CHWs were incentivised based on treatment adherence. Chang [50] found no significant differences in HIV patients'<95% pill count adherence (RR = 1.01, 95% CI 0.97–1.06, P = 0.59) or loss to follow-up (RR = 1.29, 95% CI 0.50–13.32, P = 0.60). Omer [49] found that pregnant women in intervention communities were more likely to attend at least one prenatal check-up (RR = 1.94, 95% CI 1.56-2.41, P < 0.0001), give colostrum to their newborn babies (RR = 1.21, 95% CI 1.12-1.32, P < 0.0001), and maintain exclusive breastfeeding for four months but equally likely to stop routine heavy work.[1] No quantitative data was provided for the latter two

[1] A woman advised by a CHW to reduce routine heavy work, however, was more likely to adhere to the advice than someone who was not so advised or who was advised by others (RR = 1.31, 95% CI 1.20 – 1.422, P < 0.0001).

outcomes. Winch [51] found that the proportion of malarial children treated exactly per IMCI norms was significantly higher in the intervention group than the control group (RR = 27.76, 95% CI 0.53-1441.23, P < 0.0001). DeRenzi [45B] found that the mean number of days patients were overdue at the clinic was lower in the intervention than the control group (U = 271.00, P <0.001, r = 0.50). In DeRenzi [45C], CHWs who had been in the SMS+Supervision group during DeRenzi [45B] performed worse relative to those who continued to receive SMS+Supervision (U = 68.00, P = 0.023, r = 0.340). The performance of CHWs who had been in the control group during DeRenzi [45B] improved; however, no statistically significant difference was found between the SMS+Supervision and SMS-only groups (U = 101.00, P = 0.880, r = 0.043).

Overall, these studies suggest that CHW performance improvement interventions can improve certain behavioural outcomes for patients. The body of evidence for behavioural outcomes for patients was graded moderate (downgraded for design and implementation as allocation concealment and random sequence generation were unclear in the majority of studies).

Secondary outcomes

Secondary outcomes were utilisation of health services, quality of care provided by CHWs, and CHW retention (Table 3).

Table 3 Secondary outcomes

Secondary outcome measure	Study	Intervention Type	Variables	Measure of Effect (95% CI or SE)
Utilisation of health services	Ashraf 2015 (N = specified in text)	Recruitment	Utilisation of health services	See text

For peer review only

	DeRenzi 2012A (N=11 CHWs)	Supervision	Closed referrals per CHW	SMD = 0.86 (-0.30, 2.03)
	Ginsburg 2014 (N=556 patients)	Equipment (I_a) /Training-Equipment (I_b)	Closed referrals	I_a: 43% I_b: 63% C: 53% P >0.05
	Winch 2003 (N=79 patients)	Training-Reporting-Integration	Closed referrals	RR = 0.78 (0.57, 1.07) P = 0.13
CHW quality of care	Andreoni 2016 (N = 337 CHWs)	Incentives	Closeness to vaccine distribution policy objective	SMD = - 0.32 (-0.54, -0.1)
	Ashraf 2015 (N= 298)	Recruitment	Home visits	B = 93.95 (SE = 37.19) P < 0.05
			Patients seen	B = 31.79(SE = 260.4) P >0.05
			Comm. mtgs organised	B = 17.06 (SE = 0.5.22) P < 0.01
	Ayele 1993 (N=100 CHWs)	Training-Supervision	Completion of government-set CHW activities	See text
	Bailey 1996 (N = 49)	Training	Knowledge score	Combined Intervention: 36.8% Control: 24.9% P > 0.05
	Bossuroy 2016 (N = 78 CHWs)	Incentives	TB case detection rate	See text
	Carrón 1994 (N=180 CHWs)	Training-Supervision	Quality of care score	See text
	Gautham 2015 (N= 16 CHW)	Equipment	Protocol compliance	See text
	Winch 2003 (N=102	Training-Reporting-Integration	Appropriate drug administration	RR = 1.45 (0.99, 2.13) P = 0.04

CHWs)	Dispensing correct amount of medication	RR = 1.22 (0.69, 2.18) P = 0.47

Four studies reported on utilisation; three examined the effect of CHW performance interventions on closed referrals [45A, 46, 51] and one on utilisation more generally [54]. Winch [51] found no difference in the rate of compliance with referral to the health centre among intervention or control patients (RR = 0.78, 95% 0.57 to 1.07, P = 0.13) but that patients in the intervention group were nearly four times more likely to be referred than those in the control group (RR=3.76, 95% CI 2.25 to 6.28, P <0.0001). Ginsburg [46] found similar results: while women in intervention arm B (smart phones plus patient navigation) were significantly more likely to attend for care versus women in intervention arm A (smart phones without navigation; 63% vs. 43%, P < 0.0001), adherence in the two intervention arms combined was the same (53%) as that of the control arm. Control arm CHWs interviewed fewer women than CHWs in either of the two intervention arms, however, and the proportion of study participants found to have an abnormal clinical breast exam was more than 3 times greater in the intervention arms (I_A=3.1%, I_B =3.2%) arms than the control arm (1.0%) (P < 0.0001). DeRenzi [45A], by contrast, found a large but statistically insignificant increase in the mean percentage of closed referrals in the intervention group (SMD = 0.86, 95% CI -0.30, 2.03), though a methodological limitation of the trial may explain much of the difference. Ashraf [54] found the intervention increased the number of women giving birth at the health centre by 31% (N = 1269 patients, B = 14.68, SE = 6.32, P < 0.05), the number of children under 5 undergoing health checks by 24% (N = 1618 patients, B = 318.1, se = 98.05, P < 0.05), the number of children under 5 being weighed by 23% (N = 1610 patients, B = 284.3, SE = 110.2, P < 0.05) and number of children under 1 receiving immunization against polio by 20% (N = 1530 patients, B = 14.98, SE = 4.803, P <

For peer review only

0.05). No statistically significant difference was seen in the number of postnatal (0-6 week) visits, children under 1 year receiving BCG vaccinations, and children under 1 receiving measles vaccinations.

Overall, these studies suggest that CHW performance improvement interventions can improve utilisation of services by increasing the absolute number of patients who attend the health centre and, possibly, by better identifying those who would benefit from such services. The body of evidence for utilisation of services was graded moderate (downgraded for design and implementation as allocation concealment, random sequence generation, and blinding of assessors as unclear in the majority of studies).

Eight studies reported on CHW quality of care using a range of measures [44, 47, 48, 51-55]. Four trials reported on CHWs completing [53-55] or having the knowledge necessary to complete [47] prescribed activities. Ayele [53] calculated the scores for the 13 CHW government-set activities (e.g. home visits, referrals), finding that 10 of 13 activity scores (similar at baseline, maximum score not stated, $P < 0.05$) and composite functional status score ($P < 0.0001$) were significantly higher among intervention group CHWs. Though the composite functional status score was defined as the sum of the 13 activity scores, this is no the case. The authors of the paper were contacted but were unable to supply clarifying data. Ashraf [54] found that CHWs recruited with career possibilities salient conduct 29% more household visits ($B = 93.95$, $SE = 37.19$, $P < 0.05$) and organise over twice as many community meetings (43 vs. 22, $B = 17.06$, $SE = 0.5.220$, $P < 0.01$) than those recruited with community benefits salient, while the difference in the number of patients seen at the health post is also positive but not significant ($B = 31.79$, $SE = 260.4$, $P >0.05$). Bossuroy [55] Found that, on average, the number of new TB detections increased by 2.18 (33.2%)

For peer review only

each month among intervention CHWs during the period they were incentivised based on case detection (B = 2.18, SE = 0.95, P < 0.05). The number of defaults, however, was significantly larger among intervention CHWs during this period (B = 0.08, SE = 0.04, P < 0.05). Bailey [47] found that CHWs in the intervention group demonstrated a significantly greater ability to correctly diagnose diarrhoea of varying type and severity (I = 77.3%, C = 43.1%, P < 0.05). Though non-significant differences favouring the intervention group were found in CHWs' diarrhoea knowledge, referral recommendations, and treatment practices.

Two studies reported on treatment protocol compliance [44, 51]. Winch [51] found that intervention CHWs were significantly more likely to appropriately administer drugs (RR = 1.45, 95% CI 0.99 to 2.13, P = 0.04). The study also reported that more intervention CHWs (58.9%) sold the correct amount of medication than control CHWs (48.1%). Using the ICC provided (0.210), a risk ratio was approximated (RR = 1.22, 95% CI = 0.69 to 2.18, P = 0.47). The p-value calculated by the authors (P = 0.042) indicates statistical significance; emails attempting to resolve the discrepancy received no response. Gautham [44] calculated mean protocol adherence for the intervention and control groups across a variety of stratifications of interest. The intervention group exhibited greater protocol compliance across patient order and across sexes than the control group, though the initial protocol compliance superiority of the intervention group diminished over time. Authors were contacted for the unstratified mean protocol compliance scores but were not able to provide them.

Two trials reported on observed differences in care provided [48, 52]. Carrón [52] found quality of care provided improved in all four groups (P < 0.05). The largest quality gain was observed in CHWs who received the decision tree, though

For peer review only

differences between groups were not statistically significant. Measures of variability were not reported. Andreoni [48] found that, when outliers were excluded, CHWs who receive tailored contracts were one-third closer to the policy objective than their untailored counterparts (SMD = -0.32, 95% CI -0.54 to -0.1) with an even greater effect for day-of allocation (B = 0.130, SE = 0.037, P < 0.01).

Overall, these studies suggest that CHW performance interventions can improve CHW quality of care in terms of upstream measures like completion of prescribed activities and downstream measures like adherence to treatment protocols and observed differences in the calibre of care provided. The body of evidence for CHW quality of care was graded moderate (downgraded for design and implementation as allocation concealment, random sequence generation, and blinding of assessors as unclear in the majority of studies).

No studies reported on CHW retention or adverse events. Though cost was not prespecified in the review protocol as a variable about which information would be collected, it is important in assessing the relative merits of various combinations of inputs to improve CHW performance and so it is reported in Appendix H.

DISCUSSION

Summary of Main Results

This review identified fourteen trials evaluating the effects of interventions for improving the performance of CHWs for community-based primary healthcare in LMICs. A structured summary laid out moderate quality evidence of the efficacy of CHW performance interventions in (i) improving certain behavioural outcomes for patients, (ii) improving utilisation of services by increasing the absolute number of patients who attend the health centre and, possibly, by better identifying those who would benefit from such services, and (iii) improving CHW quality of care in terms of

For peer review only

upstream measures like completion of prescribed activities and downstream measures like adherence to treatment protocols and observed differences in the calibre of care provided (e.g. technical competence). The evidence for biological outcomes was graded high, though the lack of effect on virologic failure and mortality may not be indicative of all possible biological outcomes. None of the studies included in this review reported on possible adverse effects of these interventions, either to patients, CHWs, or the health system. The heterogeneity and poor reporting of included studies precluded meta-analysis. Evidence is insufficient to draw conclusions regarding the effects of such interventions on retention or to specify which performance improvement intervention strategies are likely to be most effective. Furthermore, the nearly half of studies were compound interventions, making it difficult to isolate the effects of individual performance improvement intervention components.

That said, there is moderate quality evidence that the following practices improve behavioural outcomes for patients, utilisation of services, and/or CHW quality of care: (1) When recruiting CHWs, emphasising career possibilities rather than benefits to the community [54]. (2) When supervising CHWs, providing escalating reminders for tasks that are overdue and following-up with underperforming CHWs [45ABC]. (3) When incentivising CHWs, tailoring incentives to measure individual preferences—but only for CHWs performing a single repetitive task, not for CHWs who must perform multiple or more complex tasks [48, 55]. (4) When equipping CHWs, using mobile phone-based procedural guidance applications [44, 46]. These four strategies ought to be interpreted with caution, however, for reasons discussed in the following two sections.

Implications for research

Methodological quality of the fourteen included studies was difficult to assess

For peer review only

due to incomplete reporting of key methodological and clinical features; contacting the trialists did not always yield sufficient additional information to make such judgments. Missing data and unit of analysis problems made the size [49] and precision [51-53] of several outcomes difficult to ascertain.

The selection and validity of outcome measures is also a concern across studies. For example, several measurement instruments were employed to assess "quality of care." Some were more valid than others: the "functional status score" employed in Ayele [53], for example, measures activity, not quality—the two are not the same and should not be confused. Similarly, while Winch [51] and Omer [49] relied on patient report to assess the impact of CHWs, more rigorous, direct assessments of performance—like the service tests used by Carrón [52] and Gautham [44]—paired with biological measures would have been preferable. The lack of consensus building efforts on metrics to CHW performance improvement is a hindrance to the accumulation of knowledge in this area.

Finally, seems reasonable to consider that most of the interventions produced a considerable Hawthorne effect. A number of studies documented that existing CHWs were under-supported in their work, and the intervention is likely to have been a motivational factor [44, 45ABC]. Moreover, the novelty—and in turn the effects—of such inputs may wear off over time [45ABC]. Future studies would do well to adjust for the attention effect and assess the long-term effects of such interventions.

Implications for policy and practice

Several factors limit the applicability of these results [56]. Diverse outcomes across studies meant there was little data for each outcome of interest and multiple studies [44, 45] were underpowered to detect meaningful clinical difference. Because no studies reported adverse effects it is difficult to draw conclusions regarding the trade-off between benefits and harms [57].

For peer review only

Similarly, trials often failed to report relevant information regarding context. This is true both of the cultural context in which CHWs operate and CHW program characteristics (e.g. recruitment, incentivisation etc.) that remained unmodified in the interventions. The following factors about a potential implementation setting ought to be considered by decision-makers when assessing potential program-level performance improvement interventions [18, 57]:

1. Whether there are important differences in health system arrangements that may alter the feasibility and acceptability of the given intervention. For example, in Zambia [54] CHWs are part of the civil service and so eligible for government sponsored "in-service training." In countries where CHW programs are less well resourced it would not be possible to emphasise career advancement possibilities during recruitment.
2. Whether there are important differences in environmental conditions or other on-the-ground realities between where the studies were conducted and the implementation setting. Equipping CHWs with mobile phone-based guidance [44, 46] would not be possible in remote areas without mobile phone reception. Similarly, a supervision system premised on following-up with underperforming CHWs [45ABC], is impossible without a system to track CHW performance in the first place.
3. Whether there are important differences in the baseline health conditions between where the studies were done and the implementation setting. For example, tailoring of incentives [48] worked well for a program in which CHWs had a single categorical objective that only required one repetitive behaviour. Where CHWs have more complex objectives that require multiple behaviours performance-based incentives may distort behaviour in unintended

For peer review only

and undesirable ways [58]—e.g. in Bossuroy [55] TB defaulting increased for patients of CHWs incentivised for case detection compared with patients of CHWs receiving a fixed salary.

4. Whether studies from which evidence was drawn were conducted in settings with similar social norms. Information about, for example, whether male CHWs in the Ayele [53] study faced cultural barriers working with women in the home etc. are examples of norms and attitudes that might have shaped intervention effects and so would be relevant to interpreting trial outcomes or exploring differential effects across settings.
5. Whether sufficient resources exist to implement the proposed clinical and/or managerial support for CHWs to a high standard and maintain that standard over time. The extent to which these proposed performance improvement interventions depend on the presence of other programmatic or contextual preconditions is not clear from the included studies—e.g. one study indicated that equipment may not be as valuable in the absence of adequate training [46].

Limitations

Firstly, RCTs on CHWs remain poorly indexed in electronic databases; it may therefore be possible that, despite the very careful and extensive searches and expert consultations, some relevant RCTs were not identified. Secondly, there is no widely accepted definition for this cadre of health worker. Though we have carefully specified the inclusion criteria in keeping with previous studies, there is a need for a clear taxonomy of community health worker supports and activities. Thirdly, poor reporting and nonresponse from contacted authors meant the reviewers were unable to obtain some relevant missing data on methodological characteristics, clinical characteristics, and outcomes for some studies. Nevertheless, key data were available

For peer review only

and rigorous methods have been applied throughout, including risk of bias assessments and evaluation of the quality of the evidence.

Previous Reviews

Between the original search and the update of this review, a systematic review on intervention design factors that influence CHW performance in LMICs was published [59]. Our more sensitive search strategy and focus on trials from which causal inferences can be made, are important differences in approach.

CONCLUSION

This review concisely summarises the growing evidence base and adds to the literature by virtue of its pre-specified, systematic, and highly sensitive search strategy; inclusion of published and unpublished literature; and inclusion of only the most rigorous evidence. Mounting pressure to meet ambitious new international health goals and avoid repeating mistakes of the past, underscores the timeliness of these findings.

Contributors: MB developed the study protocol with advice from PM. MB conducted the search. MB checked eligibility criteria, conducted data extraction and drafted the manuscript. All authors commented on the manuscript.

Acknowledgements: Sincere thanks to Calla Glavin for her help with study screening and data extraction.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests: None declared.

Data sharing statement: No additional data are available.

REFERENCES

For peer review only

1. World Health Organization, *Declaration of Alma-Ata*, in *International Conference on Primary Health Care*. 1978, World Health Organization: Alma-Ata, USSR.
2. Berman, P.A., D.R. Gwatkin, and S.E. Burger, *Community-Based Health-Workers - Head-Start or False Start Towards Health for All.* Social Science & Medicine, 1987. **25**(5): p. 443-459.
3. Frankel, S., *The community health worker: effective programmes for developing countries*. 1992: Oxford University Press.
4. Heggenhougen, K., et al., *Community health workers: the Tanzanian experience*. 1987: Oxford University Press.
5. Walt, G., *CHWs: are national programmes in crisis?* Health Policy and Planning, 1988. **3**(1): p. 1-21.
6. Gilson, L., et al., *National community health worker programs: how can they be strengthened?* Journal of public health policy, 1989: p. 518-532.
7. Walt, G., *Community health workers in national programmes: just another pair of hands?* 1990, Milton Keynes, UK: Open University Press.
8. United Nations. *The Millennium Development Goals Report 2015*. 2015 [cited 2016 12 June]; Available from: http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf.
9. Nullis-Kapp, C., *Health worker shortage could derail development goals.* Bulletin of the World Health Organization, 2005. **83**(1): p. 5-6.
10. Kruk, M.E., et al., *What is a resilient health system? Lessons from Ebola.* Lancet, 2015. **385**(9980): p. 1910-1912.
11. World Health Organization. *The Addis Ababa Declaration on task shifting.* in *The first international conference on task shifting*. 2008. Addis Abbaba: Health Systems and Services, WHO.
12. Campbell, J., et al., *A universal truth: No health without a workforce: Third global forum on human resources for health, Recife, Brazil.* Global Health Workforce Alliance and World Health Organization, Geneva, 2013.
13. Moszynski, P., *One billion people are affected by global shortage of healthcare workers.* British Medical Journal, 2011. **342**: p. d696.
14. Ghani, R., *40 million children have no access to the most basic healthcare services.* British Medical Journal, 2011. **343**: p. d4667.
15. Singh, P. and J.D. Sachs, *1 million community health workers in sub-Saharan Africa by 2015.* Lancet, 2013. **382**(9889): p. 363-365.
16. Kredo, T., et al., *Task shifting from doctors to non-doctors for initiation and maintenance of antiretroviral therapy.* Cochrane Database of Systematic Reviews, 2014. **7**: p. CD007331.
17. Lassi, Z.S., B.A. Haider, and Z.A. Bhutta, *Community-based intervention packages for reducing maternal and neonatal morbidity and mortality and improving neonatal outcomes.* Cochrane Database of Systematic Reviews, 2010. **11**.
18. Lewin, S., et al., *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, 2010(3).
19. Okwundu, C.I., et al., *Home- or community-based programmes for treating malaria.* Cochrane Database of Systematic Reviews, 2013. **5**: p. CD009527.

For peer review only

20. Munos, M., et al., *Independent Evaluation of the Rapid Scale-Up Program to Reduce Under-Five Mortality in Burkina Faso.* American Journal of Tropical Medicine & Hygiene, 2016. **94**(3): p. 584-95.
21. Amouzou, A., et al., *Effects of the integrated Community Case Management of Childhood Illness Strategy on Child Mortality in Ethiopia: A Cluster Randomized Trial.* American Journal of Tropical Medicine & Hygiene, 2016. **94**(3): p. 596-604.
22. Amouzou, A., et al., *Independent Evaluation of the integrated Community Case Management of Childhood Illness Strategy in Malawi Using a National Evaluation Platform Design.* American Journal of Tropical Medicine & Hygiene, 2016. **94**(3): p. 574-83.
23. Collins, L.M., et al., *The multiphase optimization strategy for engineering effective tobacco use interventions.* Ann Behav Med, 2011. **41**(2): p. 208-26.
24. Borrelli, B., *The Assessment, Monitoring, and Enhancement of Treatment Fidelity In Public Health Clinical Trials.* J Public Health Dent, 2011. **71**(s1): p. S52-S63.
25. Hasson, H., *Systematic evaluation of implementation fidelity of complex interventions in health and social care.* Implementation Science, 2010. **5**(1): p. 1.
26. Basch, C.E., et al., *Avoiding type III errors in health education program evaluations: a case study.* Health Education & Behavior, 1985. **12**(3): p. 315-331.
27. Alam, K., J.A.M. Khan, and D.G. Walker, *Impact of dropout of female volunteer community health workers: An exploration in Dhaka urban slums.* BMC Health Services Research, 2012. **12**.
28. Olang'o, C.O., I.K. Nyamongo, and J. Aagaard-Hansen, *Staff attrition among community health workers in home-based care programmes for people living with HIV and AIDS in western Kenya.* Health Policy, 2010. **97**(2-3): p. 232-237.
29. Medhanyie, A., et al., *Knowledge and performance of the Ethiopian health extension workers on antenatal and delivery care: a cross-sectional study.* Human Resources for Health, 2012. **10**.
30. Pongvongsa, T., et al., *Determinants Of Monthly Reporting By Village Health Volunteers In A Poor Rural District Of Lao Pdr.* Southeast Asian Journal of Tropical Medicine and Public Health, 2011. **42**(5): p. 1269-1281.
31. Puchalski Ritchie, L.M., et al., *Evaluation of lay health workers' needs to effectively support anti-tuberculosis treatment adherence in Malawi.* Int J Tuberc Lung Dis, 2012. **16**(11): p. 1492-7.
32. De La Torre, S.E.C., et al., *Evaluation of community malaria worker performance in western cambodia: A quantitative and qualitative assessment*, in *61st Annual Meeting of the American Society of Tropical Medicine and Hygiene*. 2012: Atlanta, Georgia.
33. Standing, H. and A.M.R. Chowdhury, *Producing effective knowledge agents in a pluralistic environment: What future for community health workers?* Social Science & Medicine, 2008. **66**(10): p. 2096-2107.
34. World Health Organization, *Community health workers: pillars for health for all. Report of the Interregional conference, Yaoundé, Cameroon, 1-5 December 1986.* 1987: Unpublished document SHS/CIH/87.2.

For peer review only

35. Khan, K.S., et al., *Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews*. 2001: NHS Centre for Reviews and Dissemination.
36. McAuley, L. and C. Ramsay. *Cochrane Effective Practice and Organisation of Care Review Group (EPOC) Data Collection Checklist*. 2002 February 4, 2014]; Available from: https://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/datacollectionchecklist.pdf.
37. Hoffmann, T.C., et al., *Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide.* British Medical Journal, 2014. **348**.
38. Montgomery, P., et al., *The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses.* Journal of Clinical Epidemiology, 2013. **66**(8): p. 874-882.
39. Higgins, J.P.T. and S. Green, *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. 2011: The Cochrane Collaboration.
40. Guyatt, G.H., et al., *Rating quality of evidence and strength of recommendations: GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.* British Medical Journal, 2008. **336**(7650): p. 924.
41. Rehfuess, E.A. and E.A. Akl, *Current experience with applying the GRADE approach to public health interventions: an empirical study.* BMC Public Health, 2013. **13**(1): p. 9.
42. Waters, E., et al., *Essential components of public health evidence reviews: capturing intervention complexity, implementation, economics and equity.* Journal of Public Health, 2011. **33**(3): p. 462-465.
43. Movsisyan, A., G. Melendez-Torres, and P. Montgomery, *Users identified challenges in applying GRADE to complex interventions and suggested an extension to GRADE.* Journal of Clinical Epidemiology, 2015.
44. Gautham, M., M.S. Iyengar, and C.W. Johnson, *Mobile phone-based clinical guidance for rural health providers in India.* Health Informatics J, 2015. **21**(4): p. 253-66.
45. Derenzi, B., et al., *Improving Community Health Worker Performance Through Automated SMS.* ACM International Conference Proceeding Series, 2012: p. 25-34
46. Ginsburg, O.M., et al., *An mHealth model to increase clinic attendance for breast symptoms in rural Bangladesh: can bridging the digital divide help close the cancer divide?* Oncologist, 2014. **19**(2): p. 177-85.
47. Bailey, J.E. and D.W. Coombs *Effectiveness of an Indonesian model for rapid training of Guatemalan health workers in diarrhea case management*. Journal of Community Health, 1996. **21**, 269-76.
48. Andreoni, J., et al., *Using Preference Estimates To Customize Incentives: An Application To Polio Vaccination Drives In Pakistan*, in *NBER Working Paper Series*. 2016.
49. Omer, K., et al. *Evidence-based training of frontline health workers for door-to-door health promotion: a pilot randomized controlled cluster trial with Lady Health Workers in Sindh Province, Pakistan*. Patient Educ Couns, 2008. **72**, 178-85 DOI: 10.1016/j.pec.2008.02.018.

For peer review only

50. Chang, L.W., et al., *Impact of a mHealth Intervention for Peer Health Workers on AIDS Care in Rural Uganda: A Mixed Methods Evaluation of a Cluster-Randomized Trial.* AIDS and Behavior, 2011. **15**(8): p. 1776-1784.
51. Winch, P., et al., *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, 2003. **97**(5): p. 481-490.
52. Carrón, J., M. Melian, and F. Leon, *Final report of an operations research project:" Developing Tools of Low-Cost Use to Improve the Quality of Care of Rural CBD in Paraguay" Subcontract CI91. 96A January 1 1992-July 31 1994. Participating agencies: Centro Paraguayo de Estudios de Poblacion (CEPEP) the Population Council/INOPAL II.* 1994.
53. Ayele, F., A. Desta, and C. Larson, *The Functional Status Of Community-Health Agents - A Trial Of Refresher Courses And Regular Supervision.* Health Policy and Planning, 1993. **8**(4): p. 379-384.
54. Ashraf, N., O. Bandiera, and S.S. Lee, *Do-Gooders And Go-Getters: Career Incentives, Selection, And Performance In Public Service Delivery*. 2015.
55. Bossuroy, T., C. Delavallade, and V. Pons, *Fighting tuberculosis through community based counsellors: a randomized evaluation of performance based incentives in India*, in *3ie Grantee Final Report*. 2016, International Initiative for Impact Evaluation (3ie): New Delhi.
56. Schünemann, H.J., A. Fretheim, and A.D. Oxman, *Improving the use of research evidence in guideline development: 13. Applicability, transferability and adaptation.* Health research policy and systems, 2006. **4**(1): p. 25.
57. Lavis, J.N., et al., *SUPPORT Tools for evidence-informed health Policymaking (STP) 16: Using research evidence in balancing the pros and cons of policies.* Health Research Policy and Systems, 2009. **7**(Suppl 1): p. S16.
58. Benabou, R. and J. Tirole, *Intrinsic and extrinsic motivation.* The Review of Economic Studies, 2003. **70**(3): p. 489-520.
59. Kok, M.C., et al., *Which intervention design factors influence performance of community health workers in low- and middle-income countries? A systematic review.* Health Policy and Planning, 2015. **30**(9): p. 1207-27.

For peer review only

Table of Contents

Appendix A: Protocol (Finalised April 2014) 1
Appendix B: Low and Middle-Income Countries 13
Appendix C: Details on the Search Strategy 13
Appendix D: Data Extraction Form 22
Appendix E: Characteristics of Included Studies 27
Appendix F: Characteristics of Excluded, Ongoing & Studies Awaiting Classification 49
Appendix G: Risk of bias of Included Studies 55
Appendix H: Cost Data 60
Appendix I: PRISMA Checklist 64

Appendix A: Protocol (Finalised April 2014)

The following protocol was reviewed by a panel of systematic review experts from the Centre for Evidence Based Intervention, University of Oxford in April 2014.

I. PICO QUESTION

Do interventions for improving the performance of community health workers, compared with treatment as usual, result in better primary healthcare outcomes in low- and middle-income countries?

II. BACKGROUND/STATE OF THE EVIDENCE

Description of the condition

Though widely promoted as a means to provide primary healthcare in resource poor settings as early as the 1978 Alma-Ata Declaration, community health workers (CHWs) have recently become the subject of renewed interest and debate in low and middle-income countries (LMICs) facing a growing human resource crisis (Gilmore &

For peer review only

McAulife, 2013; Lehmann & Sanders, 2007, WHO, 1978). The WHO estimates the world is currently short 7.2 million healthcare workers and that this deficiency is expected to grow to 12.9 million by the year 2035 (WHO/GHWA, 2013). This phenomenon is particularly severe in sub-Saharan Africa, which bears an estimated 24% of the world's burden of illness, yet has only 3% of the health care workers and 1% of the financial resources for health care worldwide (Anyangwe & Mtonga, 2007). Most acute in rural areas, personnel shortages mean that over one billion people around the world go their entire lives without seeing a health worker (Ghani, 2011; Moszynski, 2011). This lack of human resources for health (HRH) has significantly impeded progress toward the realisation of health-related Millennium Development Goals, namely to reduce child mortality (Goal 4), improve maternal health (Goal 5), and combat HIV and AIDS, malaria and other diseases (Goal 6) (Nullis-Kapp, 2005). CHWs have been proposed as a way to fill that gap by extending services to hard-to-reach populations and remote areas. Given mounting pressure to meet Millennium Development Goals, proposals to establish lay health workers—to whom simple medical procedures can be "task shifted" from doctors—have taken on revived urgency and prominence (Singh & Sachs, 2013; WHO/UNAIDS/PEPFAR, 2008).

Systematic reviews have concluded that CHWs are effective at delivering health services as diverse as immunization, case management of acute respiratory infections, malaria prophylaxis, and preventive interventions for maternal and child health (Christopher et al., 2011; Lehman & Sanders, 2007; Lewin et al. 2010). The reviews found that CHWs had a significant impact on reducing maternal, neonatal and child mortality in resource poor settings and increased coverage of health services when

For peer review only

compared to alternative modes of service delivery organisation (Gilmore & McAulife, 2013; Haines, 2007; Lassi et al., 2010). Although the reviews are of mixed quality—not all published protocols, assessed the quality of the sources compiled, or clearly delineated the data extraction process—the results from even the most exacting among them, a Cochrane review based solely on evidence from RCTs assessed for quality, indicated that CHWs "provide promising benefits in promoting immunisation uptake and breastfeeding, improving TB treatment outcomes, and reducing child morbidity and mortality when compared to usual care" (Lewin et al. 2010, pg. 2).

While these reviews examined the effects of CHWs compared to treatment as usual or other interventions, no systematic review has yet addressed the relative effectiveness of different types of CHW interventions. The question of how to optimise CHW performance—via different approaches to recruitment, training, supervision, incentivization or other factors—remains open (Lassi et al., 2010; Swider 2002).

Why it is important to do this review

While reviews have thus far isolated principles (e.g. "competent supervision") associated with CHW performance, none have identified actionable strategies to improve performance (e.g. "weekly clinical supervision by a nurse composed of supply audits, patient checks and on-the-job training"). This is the first systematic review to consider the effectiveness of interventions to improve the performance of community health workers.

III. OBJECTIVES

For peer review only

To assess the effects of interventions to improve the performance of community health workers for community-based primary health care outcomes in low income countries.

IV. METHODS

Criteria for considering studies for this review

Types of Studies

Because health workforce outcomes are empirically testable and workplace interventions possible to randomise, randomised controlled trials, cluster-randomised trials, and factorial trials will be eligible for inclusion. Based on guidelines from the Cochrane Effective Practice and Organisation of Care Review Group (2007), if less than two controlled trials are found, data from well-designed non-randomised intervention studies such as interrupted time series with at least three time points before and after intervention, cohort, and case-control observational studies will be included.

Types of Participants

Community Health Workers:

Several definitions for and variants of the term community health worker have been employed in the literature; the roles and responsibilities of CHWs vary greatly depending on patients' access to facility-based care and the existence of other cadres of health workers. As Perry & Zulliger (2012) have noted, some CHWs have only a few days of training, while others have six months or more; some CHW cadres are paid salaries, others are volunteers; some are generalists, while others perform a narrowly defined set of interventions specific to one disease.

For peer review only

For the purpose of this review, definitions from Lewin et al. (2010) and the WHO (1987, 2007) were built on to define community health workers as any lay health workers who:

- Live and work in the community they serve
- Are part of the formal health system in a low- or middle-income country
- Perform interventions related to healthcare delivery
- Have received organised training in said interventions but have no formal or paraprofessional certification or tertiary education degree.

Although CHWs have been used in high-income countries, differences in healthcare budgets, education levels, access to care, and disease profile make it unproductive to include such actors in the comparison.

Recipients of care:

Aside from residence in a low- or middle-income country, there are no restrictions on the types of patients for whom data will be extracted.

Types of Interventions

Studies that evaluate any intervention designed to improve the performance of CHWs will be included. Studies will only be included where the primary objective, or one of the primary objectives, is to determine the efficacy or effectiveness of an intervention to improve the performance of CHWs. This includes studies like 'head-to-head' comparisons of different CHW interventions that have to date been excluded from reviews of CHW effectiveness (e.g. Lewin et al., 2010). Studies designed to examine the efficacy of a particular CHW activity (e.g. zinc tablets for diarrhoea) as opposed to the

For peer review only

relative effectiveness of different ways of supporting CHWs to deliver said activity, will be excluded. Interventions of any duration or follow up will be included.

A study will be included if the intervention applied was adequately described (i.e. described in sufficient detail to understand the key steps undertaken) in the paper, in an appendix, or by the author in a personal communication.

Types of Outcome Measures

A number of processes and outcomes might be affected by interventions that aim to improve the performance of CHWs. Reported outcomes will be extracted and categorised as follows:

- Primary:
  - Objective physical outcomes for patients
  - Subjective health outcomes for patients
- Secondary
  - Utilisation indicators (e.g. mean number of consultations/month)
  - Quality of care outcomes (e.g. accuracy of diagnosis)
  - CHW retention

Search methods for identification of studies

Electronic Searches

The following strategy was used to search MEDLINE and was adapted for other databases:

1. ((community or village) adj2 (agent$ or aid$ or promot$ or OR mobili?er$ distribut?r$ or worker$)).ti,ot,ab,kf.

For peer review only

2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliary$ or extension) adj2 (worker$ or volunteer$)).ti,ot,ab,kf.
3. (accompanier$ or accompagnateur$ or activista$ OR animatrice$ OR brigadista$ or kader$ or monitora$ or promotora$ or sevika$ or fhw$ or chw$ or lhw$ or vhw$ OR chv$ or "shastho shebika" or "shasto karmis" or anganwadi$ or "barefoot doctor" OR "agente comunitario de salud" or "agente communitario de saude").ti,ot,ab,kf.
4. (performance or effective$ or skill$).ti,ot,ab,kf.
5. 1 or 2 or 3
6. 5 and 6
7. randomized controlled trial.pt.
8. controlled clinical trial.pt.
9. randomized.ab.
10. placebo.ab.
11. clinical trials as topic.sh.
12. randomly.ab.
13. trial.ti.
14. 7 or 8 or 9 or 10 or 11 or 12 or 13
15. 6 and 14

Databases searched:

- AMED, Ovid
- BiblioMap - EPPI-Centre database of health promotion research
- British Library for Development Studies at IDS

For peer review only

- CINAHL, Ebsco
- Cochrane Central Register of Controlled Trials (CENTRAL)
- Database of Promoting Health Effectiveness Reviews (DoPHER)
- EMBASE, Ovid
- Global Health
- IDEAS database of unpublished working papers
- MEDLINE In-Process & Other Non-Indexed Citations
- MEDLINE, Ovid
- POPLINE
- Pubmed
- Social Science Citation Index/Web of Science
- WHOLIS
- World Bank's JOLIS search engine

Searching other resources

With the use of a pre-determined search strategy, studies will be identified through:

1. Systematic search of electronic databases using a pre-defined search strategy, described above.
2. Handsearch of relevant journals, search of trial registries for on-going research (e.g. ClinicalTrials.gov), and search of grey literature databases (e.g. Open GREY)
3. Review of websites and via direct contact with local agencies, research institutions, relevant government departments, and international donors and

For peer review only

multilateral agencies involved in HIV care, treatment, and prevention. These organisations will include GHWA, UNICEF, WHO, & USAID.

4. Direct contact with experts in the field for unpublished studies.
5. Review of references from included studies and relevant articles to identify other studies that met the inclusion criteria.

Data Collection and analysis

Selection of studies

Two authors will independently screen titles, abstracts, and full text articles. Included articles will be determined by consensus.

Data extraction and management

Reviewers will extract data on study design, participants, interventions, outcomes and methodological quality, using a modified version of the Cochrane EPOC Group data collection checklist. Data will be managed using Excel.

Assessment of risk of bias in included studies

The Cochrane Collaboration's tool for evaluating the risk of bias will be used to assess individual studies. The quality of evidence across a body of evidence will be assessed using the GRADE approach.

Measures of treatment effect

Statistical analysis will be carried out using the Review Manager software (RevMan 2012). Measures of treatment effect will be standardised (dichotomous: risk ratios, continuous: standardised mean difference) where possible or, if not, presented as reported by the trial authors with 95% CIs and tests of statistical significance where available.

For peer review only

Unit of analysis issues

Cluster-randomised trials will be verified to ensure appropriate analysis has been done (i.e. one that adjusts for clustering in calculating confidence intervals or P-values). If such an analysis has not been done, the necessary data (intracluster correlation coefficients - ICCs) will be extracted if possible or obtained from the investigators and results will be re-analysed. If this is not possible, point estimates, but not the reported confidence intervals or P-values, will be reported. If there are sufficiently similar studies to conduct meta-analyses, a sensitivity analysis will be undertaken using imputed ICCs based on data from the other included studies and other studies with comparable outcome measures (Campbell, 2000).

For cross-over trials, we will extract and analyse data only from the period before the crossover. For factorial studies, we will include all comparisons.

Dealing with missing data

Authors of each of the included studies will be contacted to obtain missing statistics (i.e. unreported data). Missing data due to participant attrition will be dealt with by conducting sensitivity analyses to investigate attrition as a source of heterogeneity and possible bias.

Assessment of heterogeneity

If meta-analyses is conducted, the extent of heterogeneity in results across comparable studies will be assessed using forest plots, the I^2 statistic and the Chi^2 test.

Assessment of reporting biases

Selective outcome reporting will be assessed using the approach described in Chapter 8 of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins

For peer review only

2008). Publication bias will be assessed using funnel plots (plotting effect size against standard error) to detect potential bias. Because asymmetry can be due to clinical and methodological heterogeneity in addition to publication bias, these sources of heterogeneity will also be examined as possible explanations.

Data synthesis

Given the question asked in the review is broad with respect to the type of intervention and outcome considered, a random effects analysis will be conducted if possible. If no metaanalysis is conducted, a narrative synthesis will be conducted.

Subgroup analysis and investigation of heterogeneity

No subgroup analysis will be performed.

Sensitivity analysis

A sensitivity analyses will be performed by excluding studies with a high risk of bias for any outcome for which more than one comparable study with studies with a low or moderate risk of bias.

Differences Between Protocol and Review (September 2016)

- Four additional databases were added in July of 2016 to better capture the international development literature (3ie Impact Evaluation Database, IPA Database, J-PAL Evaluations Database, World Bank Independent Evaluation Group)

References for Appendix A

Anyangwe S, Mtonga C. (2007) Inequities in the global health workforce: The greatest impediment to health in sub-Saharan Africa. *International Journal of Environmental Research and Public Health*, 4:93-100

Christopher, J. B., Le May, A., Lewin, S., & Ross, D. a. (2011). Thirty years after Alma-Ata: a systematic review of the impact of community health workers delivering curative interventions against malaria, pneumonia and diarrhoea on child mortality

For peer review only

and morbidity in sub-Saharan Africa. *Human resources for health*, *9*(1), 27. doi:10.1186/1478-4491-9-27

Ghani, R. (2011). 40 Million children have no access to the most basic healthcare services. *BMJ (Clinical research ed.)*, *343*(jul22_1), d4667. doi:10.1136/bmj.d4667

Gilmore, B., & McAuliffe, E. (2013). Effectiveness of community health workers delivering preventive interventions for maternal and child health in low- and middle-income countries: a systematic review. *BMC public health*, *13*(1), 847. doi:10.1186/1471-2458-13-847

Lassi, Z., Haider, B., & Bhutta, Z. (2010). Community-based intervention packages for reducing maternal and neonatal morbidity and mortality and improving neonatal outcomes (Review). *Public health nursing (Boston, Mass.)*, *28*(3), 246–8. doi:10.1111/j.1525-1446.2011.00953.x

Lehmann, U., & Sanders, D. (2007). *Community health workers: What do we know about them?* (pp. i–34). Geneva.

Lewin, S., Munabi-Babigumira, S., Glenton, C., Daniels, K., Bosch-Capblanch, X., VanWyk, B. (2010). Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases (Review). *Public health nursing (Boston, Mass.)*, *28*(3), 243–5. doi:10.1111/j.1525-1446.2011.00950.x

Moszynski, P. (2011) One billion people are affected by global shortage of healthcare workers. *BMJ* 2011;342:d696.

Nullis-Kapp, C. (2005). Health worker shortage could derail development goals. *Bulletin of the World Health Organization, 83*(1), 5-6.

Perry, H., & Zulliger, R. (2012). *How Effective Are Community Health Workers?* MDG Health Alliance.

Review Manager (RevMan). 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012.

Singh, P., & Sachs, J. D. (2013). 1 million community health workers in sub-Saharan Africa by 2015. *Lancet*, *382*(9889), 363–5. doi:10.1016/S0140-6736(12)62002-9

Swider, S. M. (2002). Outcome effectiveness of community health workers: an integrative literature review. *Public Health Nursing (Boston, Mass.)*, *19*(1), 11–20. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11841678

WHO (2007). Task shifting: rational redistribution of tasks among health workforce teams. Geneva: World Health Organization.

WHO/Global Health Workforce Alliance (2013) A Universal Truth: No Health Without a Workforce. Geneva [http://www.who.int/workforcealliance/knowledge/resources/hrhreport2013/en/].

WHO/UNAIDS/PEPFAR (2008) Task Shifting: Global Recommendations and Guidelines. Geneva [http://www.who.int/healthsystems/ task_shifting/en/index.html].

WHO (1978). Declaration of Alma-Ata. Paper presented at the International Conference on Primary Health Care, Alma-Ata, USSR.

WHO (1987). Community Health Workers: Working Document for the WHO Study Group. Geneva, Switzerland: WHO.

For peer review only

Appendix B: Low and Middle-Income Countries

Based on World Bank Country and Lending Groups:

Upper-middle-income economies
Albania, Algeria, Algeria, American Samoa, Angola, Azerbaijan, Belarus, Belize, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, China, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, Fiji, Gabon, Grenada, Iran, Islamic Rep., Iraq, Jamaica, Jordan, Kazakhstan, Lebanon, Libya, Macedonia, Malaysia, Maldives, Marshall Islands, Mauritius, Mexico, Mongolia, Montenegro, Namibia, Palau, Panama, Paraguay, Peru, Romania, Serbia, South Africa, St. Lucia, St. Vincent and the Grenadines, Suriname, Thailand, Tonga, Tunisia, Turkey, Turkmenistan, Tuvalu.

Lower-middle-income economies
Armenia, Bangladesh, Bhutan, Bolivia, Cabo Verde, Cameroon, Congo, Rep., Côte d'Ivoire, Djibouti, Egypt, El Salvador, Georgia, Ghana, Guatemala, Guyana, Honduras, India, Indonesia, Kenya, Kiribati, Kosovo, Kyrgyz Republic, Lao PDR, Lesotho, Mauritania, Micronesia, Moldova, Morocco, Myanmar, Nicaragua, Nigeria, Pakistan, Papua New Guinea, Philippines, Samoa, São Tomé and Principe, Senegal, Solomon Islands, Sri Lanka, Sudan, Swaziland, Syrian Arab Republic, Tajikistan, Timor-Leste, Ukraine, Uzbekistan, Vanuatu, Vietnam, West Bank and Gaza, Yemen, Zambia

Low-income economies
Afghanistan, Benin, Burkina Faso, Burundi, Cambodia, Central African Republic, Chad, Comoros, Dem. Rep Congo, Eritrea, Ethiopia, The Gambia, Guinea, Guinea-Bisau, Haiti, Dem Rep. Korea, Liberia, Madagascar, Malawi, Mali, Mozambique, Nepal, Niger, Rwanda, Sierra Leone, Somalia, South Sudan, Tanzania, Togo, Uganda, Zimbabwe

Appendix C: Details on the Search Strategy

In addition to the electronic sources listed in the main text, studies were identified through:

1. Handsearch of (a) seven journals from 1960 (or first online issue) to May 2014 (*Health Policy and Planning*, *Journal of Community Health*, *Social Science and Medicine*, *Human Resources for Health*, *BMC Public Health, Global Health Action*, and *Global Public Health*) and (b) conference proceedings from the American Society of Tropical Medicine and Hygiene.

For peer review only

2. Review of websites of nineteen leading global public health research institutions, donors, implementers, and technical agencies: International Child Health Review Collaboration; World Health Organization; United Nations Children's Fund; United Nations Development Programme; United Nations Population Fund; Joint United Nations Programme on HIV/AIDS; United States Agency for International Development; UK Department for International Human Development Resource Centre; Partners in Health; the Global Fund to Fight AIDS, Tuberculosis and Malaria; Bill & Melinda Gates Foundation; Clinton Foundation; Centre for Global Development; University Research Co-Center for Human Services (URC-CHS); CARE, Frontline Health Workers Coalition; One Million Community Health Workers Campaign; Global Health Workforce Alliance; and CHW Central.
3. Search of grey literature databases relevant to public health: Open GREY, New York Academy of Medicine Grey Literature Report, BIOSIS, and Proquest Dissertations & Theses.
4. Search of trial registries for on-going research: WHO International Clinical Trials Registry Platform, the ISRCTN Register, and ClinicalTrials.gov.
5. Direct contact with study authors and leading experts in the field of community health to solicit potentially relevant unpublished papers, on-going research, and suggestions for other contacts (see Appendix 6a for full list).
6. Examination of the reference lists of related reviews and primary studies to identify other studies that meet the inclusion criteria.

Electronic search strategies are as follows:

AMED, Ovid

For peer review only

AMED (1985 to April May 2014) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 53 hits. The search was updated May 11, 2016 and produced 4 hits.

BiblioMap: EPPI-Centre database of health promotion research

The search strategy used for BiblioMap was as follows:

1. Community Health Agent
2. Community Health Aides
3. Community health promoter
4. Community mobilizer
5. Community drug distributor
6. Community health worker
7. Village health worker
8. Rural Health Worker
9. Lay Health Worker
10. Lady health worker
11. Nutrition worker
12. Frontline health worker
13. Barangay health worker
14. Basic health worker
15. Auxiliary health worker
16. health extension worker
17. community health volunteer
18. village health volunteer
19. accompanier
20. accompagnateur
21. activista
22. animatrice
23. brigadista
24. kader
25. promotora
26. monitora
27. sevika
28. fhw
29. lhw
30. chw
31. vhw
32. chv
33. shastho shebika
34. shasto karmis
35. anganwadi
36. barefoot doctor

For peer review only

37. agente comunitario de salud
38. agente communitario de saude
39. 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
40. performance
41. effectiveness
42. skill
43. 40 OR 41 OR 42
44. 39 AND 43

The search was conducted on June 30, 2014 and produced 8 hits. The search was updated May 11, 2016 and produced 0 hits.

British Library for Development Studies at IDS

The search strategy used for IDS was as follows: "community health workers"

The search was conducted on June 30, 2014 and produced 5 hits. The search was updated May 11, 2016 and produced 4 hits.

CINAHL, EBSCO

CINAHL, Cumulative Index to Nursing and Allied Health (1960 to May 2014) was searched via Ebsco using the following terms:

> (TI ((community OR village) N2 (agent* OR aid* OR promot* OR mobili?er* OR distribut?r* OR worker)) OR TI ((village OR rural OR lay OR lady OR nutrition OR frontline OR barangay OR basic OR auxiliar* OR extension) N2 (worker* OR volunteer*)) OR TI ((accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR monitora* OR promotora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente communitario de saude"))) AND AB ((performance or effective* or skill*))

The search was conducted on May 4, 2014 and produced 189 hits. The search was updated May 11, 2016 and produced 31 hits.

For peer review only

CENTRAL

The search strategy used for the Cochrane Central Register of Controlled Trials (Issue 5 of 12, May 2014) was as follows:

1. ((community or village) near/2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)):ti,ab
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliar* or extension) near/2 (worker* or volunteer*)):ti,ab
3. (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or promotora* or monitora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "shastho shebika" or "shasto karmis" or anganwadi* or "barefoot doctor" or "agente comunitario de salud" or "agente communitario de saude"):ti,ab
4. (performance or effective* or skill*):ti,ab
5. #1 or #2 or #3
6. #4 and #5

(In: Trials)

The search was conducted on May 4, 2014 and produced 256 hits. The search was updated May 11, 2016 and produced 129 hits.

Database of Promoting Health Effectiveness Reviews (DoPHER) search strategy
The search strategy used for DoPHER was as follows:

1. Community Health Agent
2. Community Health Aides
3. Community health promoter
4. Community mobilizer
5. Community drug distributor
6. Community health worker
7. Village health worker
8. Rural Health Worker
9. Lay Health Worker
10. Lady health worker
11. nutrition worker
12. frontline health worker
13. Barangay health worker
14. basic health worker
15. Auxiliary health worker

For peer review only

16. health extension worker
17. community health volunteer
18. village health volunteer
19. accompanier
20. accompagnateur
21. activista
22. animatrice
23. brigadista
24. kader
25. promotora
26. monitora
27. sevika
28. fhw
29. lhw
30. chw
31. vhw
32. chv
33. shastho shebika
34. shasto karmis
35. anganwadi
36. barefoot doctor
37. agente comunitario de salud
38. agente communitario de saude
39. 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
40. performance
41. effectiveness
42. skill
43. 40 OR 41 OR 42
44. 39 AND 43

The search was conducted on June 30, 2014 and produced 8 hits. The search was updated May 11, 2016 and produced 0 hits.

EMBASE search strategy

EMBASE (1974 to 2014 Week 18) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 2526 hits. The search was updated May 11, 2016 and produced 796 hits.

Global Health search strategy

For peer review only

Global Health (1973 to 2014 Week 18) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 1341 hits. The search was updated May 11, 2016 and produced 450 hits.

IDEAS database of unpublished working papers

The search strategy used for the IDEAS database was as follows:

> "community health workers" performance
> Match: all; Words forms: exact; In: abstract

The search was conducted on June 30, 2014 and produced 4 hits. The search was updated May 11, 2016 and produced 4 hits.

Impact Evaluation Database (3ie)
Keyword: community health worker
Evaluation method: Randomised Control Trials (RCT)

The search was conducted on July 12, 2016 and produced 62 hits.

IPA (Innovations for Poverty Action) Database
Keyword: community
Program area: Health

The search was conducted on July 12, 2016 and produced 6 hits.

Databases within ISI Web of Science

Databases within ISI Web of Science (Science Citation Index Expanded, Social Sciences Citation Index, Conference Proceedings Citation Index – Science, Conference Proceedings Citation Index - Social Sciences & Humanities, Book Citation Index– Science, Book Citation Index– Social Sciences & Humanities) (1945-May 2014) were searched using the following strategy:

> TITLE: (((community OR village) NEAR/2 (agent* OR aid* OR promot* OR mobili*er* OR distribut*r* OR worker*))) OR TITLE: (((village OR rural OR lay

For peer review only

OR lady OR nutrition OR frontline OR barangay OR basic OR auxiliar* OR extension) NEAR/2 (worker* OR volunteer*))) OR TITLE: ((accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR promotora* OR monitora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente communitario de saude")) AND TITLE: ((performance or effective* or skill*))
Timespan: All years. Indexes**:** SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH,

The search was conducted on May 4, 2014 and produced 3047 hits. The search was updated May 11, 2016 and produced 657 hits.

J-PAL (Abdul Latif Jameel Poverty Action Lab) Evaluations Database
Keyword: Community worker
Sector: Health

The search was conducted on July 12, 2016 and produced 23 hits.

MEDLINE search strategy

MEDLINE (1946 to May 2014 Week 19) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 2143 hits. The search was updated May 11, 2016 and produced 405 hits.

MEDLINE In-Process In-Process & Other Non-Indexed Citations search strategy
MEDLINE In-Process (1946 to May 2014 Week 19) was searched via OVID using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 251 hits. The search was updated May 11, 2016 and produced 734 hits.

POPLINE, K4Health

The search strategy used for POPLINE (1965-May 2014) was as follows:

All fields: effective* OR performance OR skill AND

Title: ((community or village) near/2 (agent* or aid* or promot* or mobili*er* or

For peer review only

distribut*r* or worker*)) OR ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliar* or extension) near/2 (worker* or volunteer*)) OR (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or promotora* or monitora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "shastho shebika" or "shasto karmis" or anganwadi* or "barefoot doctor" or "agente comunitario de salud" or "agente communitario de saude")

The search was conducted on May 4, 2014 and produced 20 hits. The search was updated May 11, 2016 and produced 168 hits.

Pubmed

The search strategy used for Pubmed (1960-May 2014) was as follows:

((((("Community health agent" or "Community Health Aides" or "Community health promoter" or "Community mobilizer" or "Community drug distributor" or "community health worker" or "Village health worker"[Title/Abstract])) OR ("Rural Health Worker" or "Lay Health Worker" or "Lady health worker" or "nutrition worker" or "frontline health worker" or "Barangay health worker" or "basic health worker" or "Auxiliary health worker" or "health extension worker" or "community health volunteer" or "village health volunteer"[Title/Abstract])) OR (accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR promotora* OR monitora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente communitario de saude"[Title/Abstract])) AND (performance or effective* or skill*[Title/Abstract]))NOT MEDLINE[sb])

The search was conducted on May 4, 2014 and produced 195 hits. The search was updated May 11, 2016 and produced 233 hits.

WHOLIS

WHOLIS (1948-May 2014) was searched via the Global Health Library the following terms: "community health workers" and performance.

For peer review only

The search was conducted on June 30, 2014 and produced 3 hits. The search was updated May 11, 2016 and produced 0 hits.

World Bank's JOLIS

JOLIS was searched with the following terms:

"community health workers" and performance

The search was conducted on June 30, 2014 and produced 6 hits. The search was updated May 11, 2016 and produced 0 hits.

World Bank Independent Evaluation Group

Advanced Search>>[all items]>>Systematic Reviews and Impact Evaluations>>Health, Nutrition and Population

The search was conducted on July 12, 2016 and produced 7 hits.

Appendix D: Data Extraction Form

STUDY ID (Author name, Year):	

VARIABLE	VALUE/ JUDGE-MENT	SOURCE	OPTIONS	VARIABLE TYPE
• INCLUSION CRITERIA				
Types of Studies				
Study design			• RCT • C-RCT • Factorial • ITS • Cohort • Case Control • CBA • Other (specify)	Nominal
Types of Participants				
Live and work in the catchment they serve.			• Yes • No • Unclear	Nominal

For peer review only

Are primarily based in said community			• Yes • No • Unclear	Nominal
Belong to of the formal health system in a low-income country			• Yes • No • Unclear	Nominal
Perform interventions related to healthcare delivery.			• Yes • No • Unclear	Nominal
Have received organised training in said interventions but have no formal or paraprofessional certification or tertiary education degree			• Yes • No • Unclear	Nominal
Types of Interventions				
Primary objective to determine the efficacy or effectiveness of an intervention to improve the performance of CHWs			• Yes • No • Unclear	Nominal
Primary objective to examine the efficacy of a particular CHW activity as opposed to the relative effectiveness of different ways of supporting CHWs to deliver said activity			• Yes • No • Unclear	Nominal
Type of intervention			• Recruitment • Training • Supervision • Incentives • Other (describe)	Nominal
Types of Outcomes				

For peer review only

List				Narrative
Decision				
Include or exclude			• Include • Exclude	Dichotomous
Reason for exclusion:				Narrative
•METHODS				
Study duration			• # months	Continuous
Date of study			• MM/YYYY-MM/YYY	
Unit of allocation			• Patient • Provider • Community • Other (specify) • Unclear	Nominal
•PARTICIPANTS				
Total number			• # intervention, # control	Continuous
Setting			• Rural • Urban • Mixed (specify)	Nominal
Country				Nominal
Local name for CHWs				Nominal
Definition of CHW				Narrative
Age			• Mean (years)	Continuous
Sex			• % Male and % Female	Continuous
Education level			• Some primary • Primary • Some secondary • Secondary • Tertiary	Ordinal
Services provided				Narrative
Patients served/CHW			• Mean	Continuous
Reimbursement			• Patient pays fee for service • Patient receives care for free	Nominal

For peer review only

•INTERVENTION DESCRIPTION				
Number of intervention groups			•#	Continuous
Description of Intervention (who delivered, format and timing of delivery)				Narrative
Description of control				Narrative
Integrity of intervention				
•OUTCOMES				
List outcomes (definition, unit of measurement)				Narrative
For each outcome, time points (i) collected (ii) reported				
Unit of analysis			• Patient • Provider • Community • Other (specify) • Not clear	Nominal
Adverse or unintended effects?				Narrative
•RESULTS				
For each intervention group:				
Number of participants allocated			• #	Continuous
Sample size			• #	Continuous
Missing participants			• #	Continuous
Summary data			• 2x2 OR Mean, SD	
•PROGRAM CHARACTERISTICS				
Recruitment				

For peer review only

Selected by			• Community • Health system • Both • Other (specify)	Nominal
Process			• Describe	Narrative
Training				
Amount			• <1 day • 1 day-1 week • 1 week - 3 months • > 6 months	Ordinal
Frequency			• Once • Weekly • Monthly • Quarterly • Yearly	Ordinal
Assessment			• Required (describe) • Not required	Dichotomous
Supervision				
Type			• Clinical • Peer • None	Nominal
Frequency			• None • Daily • Weekly • Monthly • Quarterly	Ordinal
Incentivisation				
Type			• Monetary (amount) • Non-monetary • None	Nominal
•MISCELLANEOUS				
Funding source				
Power calculation			• Done • Not clear • Not done	Nominal
Conclusions of study authors				Narrative

For peer review only

Equity considerations: evidence of differential impact on different parts of the population				Narrative
•Notes				

Appendix E: Characteristics of Included Studies

Andreoni 2016

Trial reg.	AEARCTR-0000417
Country	Pakistan
Methods	STUDY DESIGN: RCT (2x2 Factorial) UNIT OF RANDOMISATION: LHW
Population	LOCAL TERM FOR CHW: Lady Health Workers (LHWs) CHWs: 100% female, mean 10 years of education, age not specified PATIENTS: Children under five HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: (A): 85 CHWs, (B): 84 CHWs CONTROL SAMPLE SIZE: (A): 88 CHWs, (B): 80 CHWs
Interventions	OBJECTIVE: Use structural estimates of time preferences to customise incentives for a sample of polio vaccinators during a series of door-to-door immunization drives in Pakistan and experimentally evaluate the effect of matching contract terms to individual discounting patterns. The study investigates the effect of tailoring, but also whether tailoring is more or less effective depending on when allocations are made. INTERVENTION DESCRIPTION: The function of LHWs is to provide oral polio vaccine to children during government organized vaccination drives; LHWs are given a supply of oral vaccine and a neighbourhood map and are asked to travel door-to-door vaccinating children with a suggested target for vaccination. LHWs are offered a fixed bonus of 1000 rupees (around $10) for completing a total of V = 300 vaccinations over a two- day drive. Vaccinators set daily targets v1 and v2 corresponding to vaccinations on day 1 and day 2 of the drive, respectively. If either of the vaccination targets, v1 or v2, are not met, the bonus is not received. Vaccinators are randomly assigned an interest rate, R, such that a single vaccination that is allocated to day 2 reduces by R the number of vaccinations required on day 1. That is, v1 and v2 satisfy the intertemporal budget constraint. There were two intervention groups: (A) Contracts tailored to individual discounting patterns with advance allocation choice and (B) Contracts tailored to individual discounting patterns with immediate allocation choice. TRAINING DURATION (DAYS): 1 TRAINING DESCRIPTION: No description of baseline training. All CHWs received a two hour training on the smartphone vaccination application and its correct use. Independent of treatment status, LHWs were also trained on the intertemporal bonus contracts and the process by which allocations were made and submitted. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Daily

For peer review only

	SUPERVISION DESCRIPTION: At the end of each day of the drive, LHWs in each neighbourhood convened at a local clinic with their supervisor, and self-reported their vaccinations for the day. Though, in principle, a monitor could investigate the neighbourhood chalk markings, in practice, this system provided for virtually no monitoring. INCENTIVE: Monetary INCENTIVE DESCRIPTION: No explicit incentives for completing vaccinations were provided and LHWs received a fixed daily wage of 100 rupees (around $1). As part of the trial, LHWs were offered fixed bonus of 1000 rupees (around $10) for completing a total of V = 300 vaccinations over a two- day drive. RATIO CHW TO POPULATION: N.S. START DATE: September 26, 2014 END DATE: December 9, 2014 CONTROL DESCRIPTION: Prior to our project, LHWs self-reported achievement and no technology existed for monitoring vaccinations. Each vaccinator in the sample was provided a smart-phone, equipped with a reporting application that permits precise observation of when and where vaccinations are conducted. There were two control groups: (A) Untailored contracts and advance allocation choice and (B) Untailored contracts and immediate allocation choice.
Outcomes	PATIENTS: None CHWs: Closeness to the policy objective DURATION OF FOLLOW UP: Immediate
Funding source	International Growth Center (IGC), National Science Foundation, grant SES-1427355, and from the Department for International Development (DFID) Building Capacity for Research Evidence (BCURE) pilot program initiative
Notes	BASELINE INPUT: Smartphones, as described above

Risk of bias

Item	**Reviewers' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "We randomly provide half of LHWs with a tailored contract and half with a random contract" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear risk	Comment: Not specified.
Blinding of participants and personnel (performance bias)	Unclear risk	Comment: Not specified.
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Not specified.
Incomplete outcome data (attrition bias)	Low risk	Comment: None reported
Selective reporting (reporting bias)	Unclear risk	Comment: Outcomes listed in the protocol were embargoed so it is not possible to make a judgement

Ashraf 2015

Trial reg.	None
Country	Zambia
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: District CLUSTERS: 24 Intervention, 24 Control

For peer review only

Population	LOCAL TERM FOR CHW: Community Health Agent (CHA) CHWs: 30% female, 18-45 years (mean 27 years), minimum 12 years of education PATIENTS: Women and children under 5 years, mean 4.2 years education HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 149 CHWs CONTROL SAMPLE SIZE: 149 CHW
Interventions	OBJECTIVE: Identify the effect of advertising career incentives at the recruitment stage. INTERVENTION DESCRIPTION: Recruitment materials make career possibilities salient by highlighting that CHAs are part of the Ministry of Health's hierarchy and that this gives them access to a career path leading to higher-ranked positions such as nurse, clinical officer, and doctor. (Being part of the civil service, CHAs are eligible for "in- service training," meaning that they attend school as a serving officer and the government pays their tuition for all of their training.) Example line from the poster: "Become a community health worker to gain skills and boost your career!" TRAINING DURATION (DAYS): 365 TRAINING DESCRIPTION: Learn basic medical practices at a residential training. REFRESHER OR ONGOING TRAINING: Refresher training referenced but not described. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: 5% of home visits are audited; CHWs meet with their supervisors at an unspecified frequency. INCENTIVE: Monetary INCENTIVE DESCRIPTION: USD 290/month RATIO CHW TO POPULATION: N.S. START DATE: August, 2010 END DATE: January, 2014 CONTROL DESCRIPTION: Recruitment materials make salient benefits to the community, thus making the CHA position look similar to existing informal positions (e.g., village health workers, traditional birth attendants, barefoot doctors) that are common in these areas. Example line from the poster: "Want to serve your community? Become a community health worker!"
Outcomes	PATIENTS: Use of health services CHWs: Number of household visits completed over the study period, number of patients seen at health post, number of community meetings organised DURATION OF FOLLOW UP: 18 months
Funding source	IGC, JPAL Governance Initiative and HBS DFRD
Notes	BASELINE INPUT: CHAs, after a year of training, are required to devote 80% of their time (4 out of 5 working days per week) to household visits. The visits' main goals are to provide advice on women's health—including family planning, pregnancy, and postpartum care—and child health—including nutrition and immunizations. In addition, CHAs are expected to inspect the household and provide advice on health-related practices such as safe water practices, household waste management, sanitation, hygiene and ventilation. During visits, CHAs are also tasked with providing basic care to any sick persons and referring them to the health post as needed. In the remaining time, CHAs are expected to assist staff at the health post (the first-level health facility in rural Zambia) by seeing patients, assisting with antenatal care, and maintaining the facility. They are also supposed to organize community meetings such as health education talks at the health post and in schools.

Risk of bias

Item	Reviewers' judgement	Support for judgement

For peer review only

Random sequence generation (selection bias)	Unclear risk	Quote: "Random assignment of the 48 districts is stratified by province and average district-level educational attainment" Comment: Unclear how this was done.
Allocation concealment (selection bias)	Unclear risk	Comment: Not specified.
Blinding of participants and personnel (performance bias)	Unclear risk	Comment: Not specified.
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Not specified.
Incomplete outcome data (attrition bias)	High risk	Quote: "18% of CHAs drop out" after one year. "This attrition rate is balanced across treatments"
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol was identified
Recruitment bias	Low risk	Comment: CHWs were recruited after the study, but by district officials exposed to one treatment and unaware of the other
Baseline imbalance	Unclear risk	Comment: Not specified
Loss of clusters	Unclear risk	Comment: Not specified
Unit of analysis issues	Low risk	Quote: "standard errors clustered at the district level"

Ayele 1993

Trial reg.	None.
Country	Ethiopia
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: District CLUSTERS: Intervention: 1, Control: 1
Population	LOCAL TERM FOR CHW: Community health agents (CHAs) CHWs: 100% Male. Mean age 31.0 years. Mean schooling 5.0 years. PATIENTS: Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 50 CHWs, number of patients unspecified CONTROL SAMPLE SIZE: 50 CHWs, number of patients unspecified
Interventions	OBJECTIVE: Determine the effect of a refresher course and monthly supervision on the health service activities of CHWs INTERVENTION DESCRIPTION: Intervention group (N= 50 CHWs) received (1) Training: refresher course (2) Supervision: monthly INTERVENTION TYPE: Training-Supervision TRAINING DURATION (DAYS): 5 TRAINING DESCRIPTION: CHWs in the intervention district received (1) a five-day refresher course based on the "expected functions of the CHWs and other relevant responsibilities which they are asked to carry out in the provision of the health care to their communities" REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: Educative supervision by health unit staff. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S.

For peer review only

	RATIO CHW TO POPULATION: N.S. START DATE: 06/1990 END DATE: 12/1990 CONTROL DESCRIPTION: CHWs (N = 52 CHWs) received nothing after baseline input (see notes). CHWs in control group had gone an average of 1.9 years since last refresher training and 30.8% had never received supervision.
Outcomes	PATIENTS: None. CHWs: (1) Composite functional status score (FSS) computed by summing the scores for the 13 CHW activities (2) 13 activity scores: (i) Outreach (days/mo.), (ii) Health education (sessions/mo.), (iii) Environmental health (activities/mo.), (iv) Maternal and child health (activities/mo.), (v) Expanded programme on immunization (activities/mo.), (vi) Births (# registered/mo.), (vii) Deaths (# registered/mo.), (viii) Curative (pts./mo.), (ix) Reports (#/mo.), (x) School health (activities/mo.), (xi) Home visits (#/mo.), (xii) Referrals (#/3mo), (xiii) Epidemic control measures (#/3mo). DURATION OF FOLLOW UP: Collected monthly for 6 months and reported at 3 months and 6 months.
Funding source	Not specified.
Notes	BASELINE INPUT: (1) CHWs were provided with registration books, standard Ministry of Health of Ethiopia monthly report forms, and stationery sufficient to last a minimum of one year. (2) General meetings were held in each peasant association with the community leaders and the community at large to discuss issues related to the CHW's job description, roles, and responsibilities. (3) Communities were encouraged and helped to establish or strengthen health committees in their respective villages.

Risk of bias

Item	**Reviewers' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "districts were randomly assigned" (p. 380) Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra supervision and refresher training).
Blinding of outcome assessment (detection bias)	Low risk	Quote: "These supervisors were aware of the study but not the district comparison or study hypothesis."
Incomplete outcome data (attrition bias)	Low risk	Comment: 0% CHW attrition (personal communication with author)
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Recruitment bias	Low risk	Comment: No individuals recruited to the health zones post randomisation.
Baseline imbalance	High risk	Comment: baseline comparability of clusters was reported, not adjusted for.
Loss of clusters	Low risk	Comment: no clusters were lost
Unit of analysis issues	High risk	Clustering not taken into account in the analysis

Bailey 1996

Trial reg.	None

For peer review only

Country	Guatemala
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Promotore
Population	LOCAL TERM FOR CHW: Promotores (Village health promoters) CHWs: N.S. PATIENTS: N.S. HEALTHCARE SETTING: Community INTERVENTION SAMPLE SIZE: 22 CHWs CONTROL SAMPLE SIZE: 27 CHWs
Interventions	OBJECTIVE: Assess the effectiveness of a newly developed method for rapid tiered training of health workers in improving community health worker knowledge and case management skills. INTERVENTION DESCRIPTION: Train CHWs using the interactive "Kader" method, developed in West Java, Indonesia. The Kader method employs tiered training whereby trainers teach trainees exactly as the trainers themselves have been taught. The training relies on role-playing by trainees using "Counselling Cards" to practice counselling or teaching, precisely as the trained village health promoters will counsel mothers in the community. The counselling cards include a diagnostic algorithm and four treatment cards outlining recommendations for five type of diarrhoea. TRAINING DURATION (DAYS): 1 TRAINING DESCRIPTION: Promotores were trained by rural health workers ("Tecnicos") in either interactively or didactically depending on trial arm REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Promotores are supervised by technicos INCENTIVE: None INCENTIVE DESCRIPTION: Promotores are volunteers RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: Train CHWs using traditional didactic training
Outcomes	PATIENTS: None. CHWs: Knowledge score. (The tests consisted of five case descriptions which were used to assess respondents' abilities to identify five diarrhoea types and the appropriate referral and treatment recommendations for each type. General knowledge regarding signs of dehydration, diarrhoea prevention, and methods for making oral rehydration solution was assessed through independent questions.) DURATION OF FOLLOW UP: Immediate
Funding source	Birmingham, Alabama chapter of The Partners of the Americas, the Cadeceus Club of the University of Alabama School of Medicine and the Larry Mayes Memorial Scholarship
Notes	BASELINE INPUT: The responsibilities of the tecnicos include: overseeing vaccination campaigns, education and supervision of promotores, and coordination of health education, promotion, sanitation and water purification efforts in their districts. The promotores are responsible for promoting Department of Public Health programs at the village level and for providing a health advice and referral network for the district health clinics.

Risk of bias

Item	**Reviewers' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "promotores were randomly assigned to receive Kader (experimental) or didactic (control) training" Comment: Unclear how this was done

Allocation concealment (selection bias)	Unclear risk	Comment: Not specified
Blinding of participants and personnel (performance bias)	High risk	Quote: "There is also possible bias secondary to the lack of blinding of participants and trainers"
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Not specified
Incomplete outcome data (attrition bias)	Low risk	Quote: "three promotores had to be excluded from analysis due to incomplete data." Comment: <6% attrition
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol was identified for this review

Bossuroy 2016

Trial reg.	None
Country	India
Methods	STUDY DESIGN: C-RCT (4 arm) UNIT OF RANDOMISATION: Health worker CLUSTERS: 78 (Intervention/Control breakdown not specified)
Population	LOCAL TERM FOR CHW: ASHA CHWs: 72% male. 50% have some university education. Recruited by Delhi-based NGO Operation ASHA PATIENTS: 57.7% male, 57.6% no school or some primary HEALTHCARE SETTING: Community/Operation ASHA centre SAMPLE SIZE: 78 (Intervention/Control breakdown not specified)
Interventions	OBJECTIVE: Investigates whether financial incentives provided to health workers may encourage them to detect new tuberculosis cases and improve treatment adherence. INTERVENTION DESCRIPTION: For the first six months, CHWs were randomly assigned to receive either a fixed salary or a salary dependent on the number of patients they had detected. In the following six months, they were randomly re-assigned to either a fixed or an incentivized salary scheme, based on the number of defaults. CHWs were randomly assigned to one of four treatment arms: they either received (A) financial incentives based on patient detection for six months and incentives based on treatment adherence subsequently, (B) financial incentives based on patient detection for six months and a fixed salary subsequently or (C) a fixed salary for six months and incentives based on treatment adherence subsequently TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: Part of intervention being tested, see above. INCENTIVE DESCRIPTION: Part of intervention being tested, see above. RATIO CHW TO POPULATION: 1:33 START DATE: December 2009 END DATE: April 2013 CONTROL DESCRIPTION: A fixed salary for the whole duration of the experiment.
Outcomes	PATIENTS: TB defaults (number of patients leaving the DOTS system during the

For peer review only

	course of their treatment) CHWs: Case detections (number of patients enrolled in the DOTS system), Health workers' effort/motivation DURATION OF FOLLOW UP: 6 months for each outcome
Funding source	3ie's donors, which include UK aid, the Bill & Melinda Gates Foundation and the Hewlett Foundation.
Notes	Note that the number of months during which each phase was carried out differs across cities, and may not be 6 months in all cases. Also during the second phase, some new CHWs who were not a part of the first phase were randomized into the experiment. BASELINE INPUT: Operation ASHA hires community-based health workers, who are each responsible for operating two DOTS centres. During an initial period of about 3 months, CHWs work on setting up the centres, getting to know the community and surroundings, and making the centre known to the local population. During the first 3 months, CHWs all receive a fixed salary. The experiment starts after the initial three months of a centre lifespan.

Risk of bias

Item	**Reviewers' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "CHWs are randomly assigned" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear risk	Comment: Unclear how this was done
Blinding of participants and personnel (performance bias)	High risk	Comment: not possible to blind receiving incentives
Blinding of outcome assessment (detection bias)	Low risk	Comment: Two primary outcomes (detection and default) were assessed using TB registers and lab registers kept by public health TB officers
Incomplete outcome data (attrition bias)	Low risk	Quote: "Even though 78 CHWs were initially randomized, 3 CHWs left Op ASHA before their baseline survey…we use all 78 CHWs for impact on attrition"
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol identified
Recruitment bias	Low risk	Comment: CHWs were working before randomisation
Baseline imbalance	Low risk	Comment: Only 5 differences out of 44 are significant at the 10 per cent level and 2 are significant at the 5 per cent level, in line with what should be expected.
Loss of clusters	Unclear risk	Comment: Not adequately described
Unit of analysis issues	Low risk	Quote: "In…all…regressions, we adjust standard errors for clustering at the health worker level since the randomization was conducted at this level."

Carrón 1994

Trial reg.	None
Country	Paraguay
Methods	STUDY DESIGN: C-RCT (2x2 Factorial) UNIT OF RANDOMISATION: Groups (each encompassing 3 districts) CLUSTERS: (A) 3 districts, (B) 3 districts, (C), 3 districts, (D) 3 districts
Population	LOCAL TERM FOR CHW: Community-based distribution (CBD) providers

For peer review only

	CHWs: 100% female. Mean age and mean education level not reported. PATIENTS: Sex, mean age and mean education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: I_a=45 CHWs, I_b=45 CHWs, I_{a+b}=45 CHWs, number of patients unspecified CONTROL SAMPLE SIZE: C=45 CHWs, number of patients unspecified
Interventions	OBJECTIVE: Develop tools of low-cost use to improve the quality of care provided by rural CHWs in Paraguay. Specifically, to determine the relative effects of (1) an interactive counselling guide based on a service algorithm and (2) a supervisory behavioural feedback instrument on the quality of CHW provided care. INTERVENTION DESCRIPTION: Group A (N = 45 CHWs) received (1) Training: use of interactive service algorithm [ABC of the Distributor]. Group B (N = 45 CHWs) received (2) Supervision: use of role-play based instrument to diagnose and provide feedback on CHW service behaviour [Behavioural Feedback Instrument]. Group C (N = 45 CHWs) received (3) Training + Supervision: Use of ABC of the Distributor service algorithm and Behavioural Feedback Instrument-based supervision. INTERVENTION TYPE: Training-Supervision TRAINING DURATION (DAYS): <1 TRAINING DESCRIPTION: (1) ABC of the Distributor: The tool was taught by nurse midwives to CHWs in 30 min individual meetings. (2) Behavioural Feedback Instrument (BFI): CHW were trained by nurse midwives in four 30 min individual meetings. REFRESHER OR ONGOING TRAINING: A two-day course in family planning service provision is imparted once a year, serving as a basic course for untrained distributors and as a refresher course for trained ones. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 07/1992 END DATE: 07/1993 KEY CHW COMPETENCY: CONTROL DESCRIPTION: Group D (N = 45 CHWs) CHWs received the usual two-day annual family planning course and yearly refresher training.
Outcomes	PATIENTS: None. CHWs: Pretest-posttest gains in quality of care provided. Assessed using (1) an observational service test in which a trained simulated client asks services of a CHW and enacts a learnt client profile as she interacts with the provider. When the consultation ends, the simulated client fills out a checklist (range: 41-53 items, depending on client profile) indicating whether the promoter emitted each of the expected provider behaviours; the quality of care score is the sum of the items checked. The service test was administered to all CHWs. (2) A client questionnaire examining 11 indicators across 5 domains: (i) Method choice (user chooses), (ii) Information given to the client (user knowledge of other methods, user knowledge of action upon forgetting, provider advised side effects/method shift), (iii) Interpersonal relations (positive user perceptions), (iv) Provider Technical Competence (physical complaints, contraindications, problems with method, satisfaction with method, unwanted pregnancies), (v) Follow-up mechanisms (provider focus on revisit). The client questionnaire was administered to one patient per CHW randomly selected from the client register. Both the service test and the client questionnaire were conducted by the same trained simulated client. DURATION OF FOLLOW UP: All outcomes collected and reported at baseline and after 12 months.

For peer review only

Funding source	US Agency for International Development (USAID), Office of Population

Risk of bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: groups were "randomly assigned" (p. 10) Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra training or supervision).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: insufficient information to permit judgement
Incomplete outcome data (attrition bias)	High risk	Comment: 156/180 CBD found, 127/180 clients found. Quote: "Provider mortality affected unevenly the distribution of the six client profiles of the Service Test across provider groups. To avoid possible distortions in group comparisons, the equal distribution of client profiles across groups was re-established through a process whereby, for each client profile, (a) the lowest number of cases was determined and (b) the number of cases in the other groups was reduced to this level through random elimination of cases. 124 cases, 31 per group remained in the analysis" (p. 13) Comment: Randomly deleting data is not the same as never having collected the data in the first place; CHWs missing from one group as a result of death may be systematically different than the ones who were randomly deleted from the other groups.
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Recruitment bias	Low risk	Comment: No individuals recruited to groups post randomisation.
Baseline imbalance	Low risk	Comment: Baseline comparability of clusters not reported, pair-matched randomisation of clusters used.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	High risk	Comment: Clustering not taken into account in the analysis.

Chang 2011

Trial reg.	NCT00675389
Country	Uganda
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Clinic catchment area CLUSTERS:
Population	LOCAL TERM FOR CHW: Community-based peer health workers (PHW) CHWs: 100% female. Mean age and mean education level not reported. PATIENTS: HIV+ patients 66% female, 34% male. Median age 34 years (range: 15-67) in intervention group and 36 (16-76) in control group.

For peer review only

	Mean education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 13 CHWs, 446 patients CONTROL SAMPLE SIZE: 16 CHWs, 524 patients
Interventions	OBJECTIVE: Determine the effect of an mHealth (mobile phone) CHW support intervention on AIDS care in rural Uganda. INTERVENTION DESCRIPTION: CHWs randomized to the intervention were each given a mobile phone, a one-day residential training, and an hour-long field-based practicum training on the mHealth intervention. These CHWs were asked to send a text message reporting adherence and clinical data back to a centralized database immediately after or during each home visit. These texts were numeric codes produced by converting all quantitative data on the home visit form to simple integers arranged in a predetermined order separated by asterisks, (e.g. pills given, taken, defaulted might convert to *282828*). CHWs in the intervention arm were also encouraged to call a Rakai Health Sciences Program mobile phone hotline or toll-free warmline (staffed only during clinic hours) with questions or concerns. Clinic staff receiving CHW texts and calls could opt to provide care instructions to CHWs, send a higher-level care provider to the patient, or arrange transport to health care facilities. TRAINING DURATION (DAYS): 2 TRAINING DESCRIPTION: PHWs received a residential training on basic HIV pathogenesis, prevention, treatment, adherence counselling, performing pill counts, protecting patient confidentiality, and filling out a home visit form. In addition to their baseline training, PHWs randomised to the mHealth Arm were each given a mobile phone, a one day residential training, and an hour long field-based practicum training on the mHealth intervention. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: PHW day-to-day supervision was largely performed by a single staff member working part-time. INCENTIVE: Financial INCENTIVE DESCRIPTION: Monthly allowance of about 12.5 USD RATIO CHW TO POPULATION: 1:34 START DATE: 04/2006 END DATE: 06/2008 CONTROL DESCRIPTION: CHWs who did not receive the mHealth intervention continued providing regular care (see baseline input).
Outcomes	PATIENTS: Primary: cumulative risk of virologic failure (any failure during follow-up period equalling failure with viral loads conducted every 24 weeks; failure defined as .400 copies/mL). Secondary: (i) patient adherence (pill counts and 3 day self-report), (ii) virologic failure at 24 and 48 weeks of ART (failure defined as 400 copies/ml), (iii) lost to follow- up (lack of a pharmacy refill visit in over 90 days), and (iv) mortality (ascertained through verbal autopsies) DURATION OF FOLLOW UP: A summary clinic pill count was calculated by dividing the number of pills taken over the study period by the sum of pills expected to be taken over the study period. Viral loads were measured at 24 to 192 week time points from antiretroviral therapy initiation by 24-week intervals. CHWs: None.
Funding source	Doris Duke Charitable Foundation, The Division of Intramural Research, The National Institute for Allergy and Infectious Diseases, National

For peer review only

	Institutes of Health, and National Institutes of Health Training (2T32-AI07291) and Career Development (1K23MH086338-01A2) Grants.
Notes	BASELINE INPUT: CHWs received a two-day residential training on basic HIV pathogenesis, prevention, treatment, adherence counselling, performing pill counts, protecting patient confidentiality, and filling out a home visit form. CHWs provided clinical and adherence monitoring and psychosocial support to fellow patients. At home visits, CHWs recorded a review of symptoms, an adherence self-report, and pill count results using a standardized form. CHWs were each given a bicycle, t-shirts, basic supplies, and an initial monthly allowance of about 12.5 USD. CHW day-to-day supervision was largely performed by a single RHSP staff member working part-time.

Risk of bias

Item	**Reviewers' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "29 PHWs at 10 clinics were randomized by clinic to receive the intervention or not." Comment: Unclear how done.
Allocation concealment (selection bias)	Low risk	Quote (parent trial): We assigned clusters using unmatched, unrestricted random allocation by a drawing of lots. Comment: Author provided the following information via email: "allocation was done simultaneously with the randomisation process"
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra training and a mobile phone).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Author provided the following information via email: "this was a biological sample…no special measures taken"
Incomplete outcome data (attrition bias)	Low risk	Quote (parent trial): 9.3% died; 2.2 % LTFU. "Efficacy analyses for cumulative risk of virologic failure and for virologic outcomes at specific time points from ART initiation were by intention to treat using log-binomial regression with generalized estimating equations (GEE)"
Selective reporting (reporting bias)	Low risk	Comment: Parent trial registered on ClinicalTrials.gov (NCT00675389). All outcomes reported.
Recruitment bias	Low risk	Comment: No individuals recruited to clinic catchment area post randomisation.
Baseline imbalance	Low risk	Quote: Sociodemographic characteristics, immunologic and clinical stage of disease, the proportion of patients on ART, the median duration of time patients were on ART prior to the start of the trial, and the pre-trial 24 week and 48 week virologic failure" were well balanced between arms.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	Low risk	Quote: "…robust variance estimates appropriate for cluster-randomised trials"

For peer review only

DeRenzi 2012A

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: RCT (Pilot study) UNIT OF RANDOMISATION: CHW
Participants	LOCAL TERM FOR CHW: community health worker CHWs: 100% Female. Age range 23-55 years. Mean schooling unreported. PATIENTS: Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 8 CHWs, number of patients unspecified CONTROL SAMPLE SIZE: 7 CHWs, number of patients unspecified
Interventions	OBJECTIVE: Evaluate the impact of SMS reminders to improve the promptness of routine CHW visits. INTERVENTION DESCRIPTION: CHWs received an SMS message two days after they had reported a referral with CommCare, reminding them of the follow-up visit. Daily SMS messages were sent until the follow-up was recorded. INTERVENTION TYPE: Supervision TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: Non-financial INCENTIVE DESCRIPTION: CHWs are incentivized with personal use of the mobile phone they are given for CommCare. RATIO CHW TO POPULATION: N.S. START DATE: 04/2010 END DATE: 06/2010 CONTROL DESCRIPTION: CHWs received nothing.
Outcomes	PATIENTS: Change in percentage of closed referrals CHWs: None TIMEPOINTS COLLECTED/REPORTED: Referral and follow-up information was collected automatically by CHWs using CommCare, and sent immediately to the server by the CHW. 24 days
Funding source	Not specified.
Notes	The study period was split into baseline and intervention periods that were 39 and 24 days long, respectively. Two CHWs in the control group were excluded from analysis because they did not report any referrals during the intervention period.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: " participants were randomly assigned" (p. 27) Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra SMS notifications)
Blinding of outcome assessment (detection	Unclear risk	Comment: insufficient information to permit judgement.

For peer review only

bias)		
Incomplete outcome data (attrition bias)	Low risk	Comment: "Two CHWs in the control group were excluded from analysis because they did not report any referrals during the intervention period."
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement

DeRenzi 2012B

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: CHW (stratified by baseline performance)
Participants	LOCAL TERM FOR CHW: community health worker CHWs: Male and Female (proportions unspecified. Age range 23-63 years. Mean schooling unreported. PATIENTS: Chronic care patients (most are HIV-positive, though some have diabetes, tuberculosis, or other long-term and chronic ailments). Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION AND CONTROL SAMPLE SIZE: 87 CHWs (breakdown unspecified, after attrition: I =34, C = 40), number of patients unspecified
Interventions	OBJECTIVE: Evaluate the impact of SMS reminders to improve the promptness of routine CHW visits. INTERVENTION DESCRIPTION: Escalating reminder system that sent SMS reminders directly to the CHW before notifying the CHW's supervisor after several overdue days (augments existing supervision structures). INTERVENTION TYPE: Supervision TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: Non-financial INCENTIVE DESCRIPTION: CHWs are incentivized with personal use of the mobile phone they are given for CommCare. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: CHWs received nothing after baseline input.
Outcomes	PATIENTS: Mean time chronic care patients overdue. CHWs: None. TIMEPOINTS COLLECTED/REPORTED: Data collected on a rolling basis, sent from CHWs post closed referral (40 days)
Funding source	Not specified.
Notes	The study period was split into baseline and intervention periods that were 40 days each. BASELINE INPUT: Pathfinder CHWs are all volunteers; they are provided with a "generous transportation stipend" during their monthly meetings and incentivized with personal use of the mobile phone they are given for CommCare. CHANGES FROM PILOT: In the pilot study (DeRenzi 2012A), the earliest reminder was sent on the evening that the CHW's visit was due, too late to make an on-time visit. For the Dar es Salaam studies (DeRenzi 2012B, DeRenzi 2012C), *proactive* reminders were incorporated, with the first one being sent the day before a visit is due. Second, if a referral was not closed during the pilot, the system continued to send SMS messages but had no means of determining *why*

For peer review only

	the CHW was not reporting a follow-up. This issue directly inspired the escalation to supervisor evaluated in DeRenzi 2012B.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: " CHWs in each bin were then randomly assigned to the intervention or control group." Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra SMS messages and supervisor intervention).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: insufficient information to permit judgement.
Incomplete outcome data (attrition bias)	High risk	Quote: "Although 87 CHWs participated in this study, 13 participants were removed before analysis for one of the following reasons: no data submitted during the intervention period (11 CHWs), no active cases (1 CHW), and technical issues with CommCare or the phone (1 CHW). This removal left 34 CHWs in the SMS+Supervisor group and 40 CHWs in the control group." "During the course of the intervention, there were occasional phone problems (e.g. accidentally deleting the CommCare application) and discrepancies (e.g. a CHW reporting s/he had already sent data) reported from CHWs in the field. If there was a discrepancy, we removed the client from our analysis. This happened for 27 clients during the intervention period, which represented 3.7% of the total number of clients visited during that period."
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement

DeRenzi 2012C

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: RCT UNIT OF ALLOCATION: CHW
Participants	LOCAL TERM FOR CHW: Community health worker CHWs: Male and Female (proportions unspecified. Age range 23-63 years. Mean schooling unreported. PATIENTS: Chronic care patients (most are HIV-positive, though some have diabetes, tuberculosis, or other long-term and chronic ailments). Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION AND CONTROL SAMPLE SIZE: 87 CHWs (breakdown unspecified), number of patients unspecified
Interventions	OBJECTVE: Although only 6.5% of cases in DeRenzi 2012B required the involvement of the supervisor, this escalation step is still more costly and burdensome than sending SMS messages alone. The objective of this study was to

For peer review only

	quantify the effect of escalating to the supervisor versus SMS messages alone in the context of SMS reminders to improve the promptness of routine CHW visits. INTERVENTION: Intervention group (SMS-Only) was provided only with proactive and overdue reminders, but escalation to the supervisor never occurred. MODE OF DELIVERY: The system, implemented on top of CommCare, first sends proactive reminders to intervention CHWs (the day before (x1) and the day of (x2) a scheduled routine visit. Daily reminders are sent while the visit remains overdue, informing the CHW od the # of days the visit is overdue. INTERVENTION TYPE: Supervision TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: Non-financial INCENTIVE DESCRIPTION: CHWs are incentivized with personal use of the mobile phone they are given for CommCare. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL: CHWs (SMS+ Supervisor) received exactly the same input as the intervention group in DeRenzi 2012B.
Outcomes	PATIENTS: None. CHWs: Mean time overdue. TIMEPOINTS COLLECTED/REPORTED: 90 days
Funding source	Not specified.
Notes	90-day baseline period and a 90-day intervention period. BASELINE INPUT: Since the baseline period followed directly after DeRenzi 2012B and with the same CHWs, some CHWs continued to receive escalating SMS reminders during the baseline, while others did not. H1: For CHWs who received escalating SMS reminders during the baseline period, the SMS-Only intervention (removing the supervisor) would decrease performance compared to continuing in the SMS+Supervisor condition. H2: For CHWs who did *not* any receive SMS reminders during the baseline period; the SMS+Supervisor intervention would result in a larger increase in performance than the SMS-Only intervention. The studies presented here measure reported follow-ups and are not correlated with ground truth of actual follow-ups.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: " CHWs were then randomly assigned to one of two conditions" Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received SMS and/or additional supervision).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: insufficient information to permit judgement.

For peer review only

Incomplete outcome data (attrition bias)	High risk	Quote: "Of the original 87 CHWs, 26 were excluded from analysis for the following reasons: incorrect phone number (2 CHWs), having no data during baseline, intervention, or both periods (22 CHWs), and technical problems with the phone (2 CHWs). In the end, there were 32 CHWs originally from the intervention group in Study 1 (16 SMS+Supervisor, 16 SMS-Only) and 29 from the control group in Study 1 (14 SMS+Supervisor, 15 SMS-Only)."
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement

Gautham 2015

Trial reg.	N.S.
Country	India
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: CHW
Population	LOCAL TERM FOR CHW: Rural health provider CHWS: 8 male, 8 female; existing rural health providers; Males: mean age of 48.75 years, 10 years of education or more. Females: mean age 24.75 years, 10 years of education or more PATIENTS: Paediatric and Adult patients HEALTHCARE SETTING: Community (personal communication with S. Iyengar) INTERVENTION SAMPLE SIZE: 8 CONTROL SAMPLE SIZE: 8
Intervention	OBJECTIVE: (a) Use a randomized control design to measure changes in protocol compliance (PC) by health workers in their everyday work settings, (b) assess the usability and acceptability of the mobile application with health workers in the experimental group and (c) obtain patient feedback on health workers' use of the mobile system during treatment. INTERVENTION DESCRIPTION: Mobile phone–based, media-rich procedural guidance applications TRAINING DURATION (DAYS): 2 days TRAINING DESCRIPTION: Training programme in guideline-based care and in use of the mHealth system by physicians who spoke the local language. The study team assessed all RHPs on their understanding of guideline-based care using a simple assessment instrument developed by the research team, on which every RHP received a pre-and post-training test score (PTTS). REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: A set of paper guidelines to use in the field
Outcomes	PATIENTS: None

For peer review only

	CHWs: Protocol compliance DURATION OF FOLLOW UP: 2 months
Funding	Information Society Innovation Fund (ISIF)

Risk of Bias

Item	Author's judgement	Description
Random sequence generation (selection bias)	Unclear risk	Quote: "At the end of the training, we randomly assigned half the RHPs to the control group and other half to an experimental group." Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear risk	Quote: "At the end of the training, we randomly assigned half the RHPs to the control group and other half to an experimental group." Comment: Unclear how this was done
Blinding of participants and personnel (performance bias)	High risk	Quote: "after the training, the RHPs in the experimental group went back to their field sites with the applications on their mobile phones, while for the ones in the control group, the data card with the application was taken out" Comment: Unclear if patients were blinded. CHWs were not blinded.
Blinding of outcome assessment (detection bias)	High risk	Quote: "PC was evaluated by direct observations of the RHPs' performance as they were treating the patients."
Incomplete outcome data (attrition bias)	Unclear risk	Comment: Not described
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol identified

Ginsburg 2014

Trial reg.	N.S.
Country	Bangladesh
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: CHW
Population	LOCAL TERM FOR CHW: Community health workers CHWS: Female; age N.S.; high school education; recruited from rural villages by study team PATIENTS: Adult women aged 25 and older and living in the catchment area of Khulna Division, Bangladesh. Exclusion criteria were women younger than age 25 and women unable to give verbal consent. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: A: 246 B: 240 CONTROL SAMPLE SIZE: C: 70
Intervention	OBJECTIVE: Test whether CHWs guided by smart phone applications would be more efficient and effective than CHWs without phones at encouraging women with an abnormal breast exam to adhere to advice regarding a clinic appointment INTERVENTION DESCRIPTION: Arm A: smart phone with applications to guide interview, report data, show motivational video, and offer an appointment for women with an abnormal CBE. Arm B: smart phone platform as in arm A plus additional "patient navigation training."

For peer review only

	TRAINING DURATION (DAYS): 3 TRAINING DESCRIPTION: CHWs and supervisors attended an intensive training program on breast anatomy, physiology, and pathology, as well as the 2-hour training session on CBEs. CHWs were taught in a gentle, culturally sensitive manner how to obtain informed verbal consent, to ask the structured interview questions, and to perform a CBE. Only the 10 CHWs randomized to arm B were given an extra day of training in patient navigation. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Three female CHW supervisors, also local women but with postsecondary education and prior experience in the health or social services sector, were hired to oversee daily workflow and to liaise with our field operations team. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:19 START DATE: July, 2012 END DATE: October, 2012 CONTROL DESCRIPTION: No smart phone; same interview recorded on paper
Outcomes	PATIENTS: The "adherence" (to advice regarding a clinic appointment) for women with an abnormal CBE CHWs: N.S. DURATION OF FOLLOW UP: 1 month
Funding	Rising Stars in Global Health award from Grand Challenges Canada.

Risk of Bias

Item	Author's judgement	Description
Random sequence generation (selection bias)	Low Risk	Quote: "The 30 CHWs were randomized by manual lottery (by O.M.G., eyes closed, hand in jar with identical folded papers with CHW names) to one of the three study arms."
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding of participants and personnel (performance bias)	Unclear	Comment: Unclear if patients were blinded. CHWs could not be blinded.
Blinding of outcome assessment (detection bias)	Low risk	Comment: clinic records used, so low risk of bias
Incomplete outcome data (attrition bias)	Unclear risk	Comment: No attrition reported
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol identified

Omer 1998

Trial reg.	None
Country	Pakistan
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Enumeration area CLUSTERS: Intervention: 5, Control: 5
Population	LOCAL TERM FOR CHW: Lady health workers (LHWs)

For peer review only

	CHWs: 100% female. Mean age and education level not reported. PATIENTS: 100% women, of whom 45% report receiving "some formal education." Mean age not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 52 CHWs, 529 patients CONTROL SAMPLE SIZE: Unspecified number of CHWs, 541 patients
Interventions	OBJECTIVE: Demonstrate the effective use of community-based evidence for health promotion by CHWs in Sindh, Pakistan. INTERVENTION DESCRIPTION: In a consultation forum convened for the purpose, CHWs and government health officials developed a targeted tool to guide household discussions. It was designed to be "user-friendly, indigenous, and attractive" to both CHWs and their clients. The tool was based on ajrak, a traditional Sindh fabric, with the message in embroidery. The design team also consulted with the provincial team of the National Programme, CHWs, women in different urban and rural communities, local designers and handicraft experts. The embroidery depicted maternal practices like attending and not attending antenatal check-ups, giving colostrum after birth and not doing heavy work. The team pre-tested the embroidery in five communities not included in this study, ensuring input from CHWs and women in the communities to adjust the tools. In intervention communities, CHWs showed these panels to pregnant women during their household visits and invited them to describe what was going on in the pictures. They encouraged their clients to express their views of the situations depicted on the cloths and discussed any difficulties they might face. To go for prenatal check-ups, for example, the discussion might include when and how the person would go. Training in the tool included an explanation of the tools, role modelling and practice in the community. INTERVENTION TYPE: Training/integration TRAINING DURATION (DAYS): Part of intervention TRAINING DESCRIPTION: Part of the intervention. REFRESHER OR ONGOING TRAINING: Part of the intervention. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Monitoring relied on the existing LHW supervisory structure in both intervention and control communities. As part of routine control visits, supervisors did random checks to see if LHWs used the tools in intervention communities. LHSs held monthly meetings at district level, where they would get feedback on progress and problems they were not able to rectify on their own. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 06/2000 END DATE: 04/2001 CONTROL DESCRIPTION: CHWs (N = unspecified) in control communities did not use the embroidered cloths and received no training in how to use them, but continued their usual work visiting women before and after childbirth to discuss safe practices in pregnancy based on the standard training of the CHW programme. (N = 541 patients)
Outcomes	PATIENTS: (i) attendance at prenatal checkups; (ii) stopping of routine heavy work during pregnancy (iii) giving colostrum to newborn babies (iv) maintaining exclusive breastfeeding for four months. CHWs: Performance (# of visits to pregnant women) DURATION OF FOLLOW UP: All outcomes were collected and reported at baseline and in a household survey after CHWs in intervention communities had used the tools for 10 months.
Funding source	Canadian International Development Agency's (CIDA) Canada Fund for Local

For peer review only

	Initiatives in Pakistan.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "We allocated these randomly using a computerised random number generator" (p. 180)
Allocation concealment (selection bias)	Unclear risk	Quote: "...using a computerised random number generator" (p. 180) Comment: Unclear if enumeration areas were assessed before number appeared.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra training on a novel communication tool).
Blinding of outcome assessment (detection bias)	Low risk	"Fieldworkers, who did not know which communities were intervention and which were control, collected data" (p. 180)
Incomplete outcome data (attrition bias)	Low risk	Quote: "Analysis followed the intention to treat principle, with all women in the intervention communities analysed as though exposed, whether or not they had seen the intervention materials." (p. 180)
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Other bias	Low risk	In cluster-randomized trials, particular biases to consider include: (i) Recruitment bias Comment: Low risk; no individuals recruited to enumeration area post randomisation. (ii) Baseline imbalance Comment: Unclear risk; baseline comparability of clusters not reported, not adjusted for. (iii) Loss of clusters Comment: Low risk; no clusters were lost (iv) Unit of analysis issues Comment: Unclear risk; not clear if clustering taken into account in the analysis. (v) Comparability with individually randomized trials Comment: Irrelevant, all studies in this review are C-RCTs.
Recruitment bias	Low risk	Comment: No individuals recruited to enumeration area post randomisation.
Baseline imbalance	Unclear risk	Comment: Baseline comparability of clusters not reported, not adjusted for.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	Unclear risk	Comment: Not clear if clustering taken into account in the analysis.

Winch 2003

Trial reg.	None
Country	Mali
Methods	STUDY DESIGN: Pair matched C-RCT UNIT OF RANDOMISATION: Health zones CLUSTERS: Intervention: 5; Control: 5

For peer review only

	DURATION: 15 months (Visual aids: 11/2001-02/2002; Referral mechanism: 03/2001-06/2002)
Population	LOCAL TERM FOR CHW: Gérant (village drug kit managers) CHWs: Sex, age and education level not reported. PATIENTS: The intervention was directed at children <14 years. Mean schooling of carers was 0.85 years. Sex of all patients treated was not reported. HEALTHCARE SETTING: Household INTERVENTION AND CONTROL SAMPLE SIZE: ~102 CHWs in both arms, 152 patients in the intervention arm, 134 in the control arm
Interventions	OBJECTIVE: Evaluate an intervention to (i) improve the skills of CHWs to counsel parents on correct home administration of chloroquine (CQ), and (ii) increase the referral of sick children to community health centres (CHC). INTERVENTION DESCRIPTION: (1) Use of visual aids for drug administration and danger sign recognition (2) Introduction of a referral/counter referral recordkeeping mechanism (3) Refresher training (4) Community engagement INTERVENTION TYPE: Training-Reporting- Integration TRAINING DURATION (DAYS): Part of the intervention TRAINING DESCRIPTION: Part of the intervention REFRESHER OR ONGOING TRAINING: a 5 day additional training course was conducted in a local elementary school that covered: (i) doses of CQ for different age groups (also covered in the standard course), (ii) how to counsel carers of young children regarding the correct administration of CQ in the home and the importance of going to the CHC when referred, (iii) indications for referral including danger signs and failure of the child's symptoms to respond to treatment, (iv) basic evaluation of a sick child for danger signs such as dehydration requiring referral, and (v) how to use the referral notebook and the visual aids SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: ~1:3 START DATE: 03/2001 END DATE: 06/2002 CONTROL DESCRIPTION: When a new village drug kit is established, all managers go through standard training which consists of a 35-d literacy course in which they are taught to read and write the national language, Bambara, using a phonetic alphabet, followed by a 1-week course on how to manage the drug kit including the correct doses to recommend and sell for each product in the kit. All of the drug kit managers in both the intervention and comparison groups had had this standard training prior to the study.
Outcomes	PATIENTS: The proportion of children with indications for referral to a health facility who were referred to and subsequently reached the facility. CHWs: (i) The proportion of children who were treated with a full course of CQ in the home DURATION OF FOLLOW UP: At least twice a week an interviewer visited the CHW, talked to him/her and examined the register of sales maintained, noting down the cases where CQ had been sold to treat a child aged <6 years. The interviewer assed the effect of the intervention on the quality and content of the counselling received from the CHW through direct questions during follow-up in the home 5 days after the sale. Data from health centre registers and CHW referral notebooks was collected at the end of the trial. All data was reported at last observation.
Funding source	United States Agency for International Development through the Family Health and Child Survival Cooperative Agreement with Johns Hopkins Bloomberg School of Public Health

For peer review only

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "1 health zone from each pair was randomly assigned to receive the intervention." Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra materials, refresher training and referral recording materials).
Blinding of outcome assessment (detection bias)	Unclear risk	Quote: "Fifteen field interviewers were recruited. In 6 of the 10 health zones there was 1 interviewer only, while the other 4 health zones had 2 or 3 interviewers due to their large size. Each interviewer was assigned a set of village drug kits and their clients to study. At least twice a week the interviewer visited the manager of the village drug kit, talked to him/her and examined the register of sales maintained by him/her, noting down the cases where CQ had been sold to treat a child aged <6 years. On the fifth day after the sale, the field interviewer went to the child's house and interviewed the person who had administered the treatment to the child. " (p. 483)□□Comment: Insufficient information to permit judgement of 'Low risk' or 'High risk'
Incomplete outcome data (attrition bias)	Unclear risk	Comment: Insufficient reporting of attrition/exclusions to permit judgement.
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Recruitment bias	Low risk	Comment: No individuals recruited to health zones post randomisation.
Baseline imbalance	Low risk	Comment: Baseline comparability of clusters not reported, pair-matched randomisation of clusters used.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	Low risk	Comment: Intracluster correlation coefficient taken into account in the analysis.

Appendix F: Characteristics of Excluded, Ongoing & Studies Awaiting Classification

Excluded studies:

Record name	Year	Reason for exclusion
Afenyadu	2005	Did not meet the definition for CHW (CHWs were teachers)
Ahmed	1993	Not a randomised controlled trial

For peer review only

Akogun	2001	Not a randomised controlled trial
Alam	2012	Not a randomised controlled trial
Amouzou	2014	Non-CHW comparator
Ashraf	2014	Did not meet the definition for CHW (CHWs were hair stylists)
Avula	2011	Not a randomised controlled trial
Baqui	2008	Non CHW comparator
Begum	2009	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Bertrand	1986	Non CHW comparator
Bhatt	1981	Not a randomised controlled trial
Bhutta	2008	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Björkman	2009	Health workers were primarily facility-based
Bojang	2011	Non CHW comparator
Boone	2008	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Brieger	2000	Health workers were primarily facility-based
Chin-Quee	2016	Not a randomised controlled trial
Curtale	1995	Not a randomised controlled trial
Das	2014	Effects of the CHW performance intervention cannot be disaggregated
Davies-Adetugbo	1997	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Elder	1992	Not a randomised controlled trial
Findley	2013	Not a randomised controlled trial
Florez-Arango	2011	Did not meet the definition for CHW (one quarter of assessed health workers were primarily facility-based)
Foreit	1984	Health workers were primarily facility-based
George	2012	Did not meet the definition for CHW (Not part of health system)
Harvey	2008	Not a randomised controlled trial
Iyengar	2013	One quarter of assessed health workers were primarily facility-based
Jintaganont	1988	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Joos	2016	Performance interventions in both arms
Jung	2009	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance

For peer review only

Kalyango	2012	Intervention not designed to improve performance
Morris	2015	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Owais	2011	Tested the efficacy of an intervention for mothers rather than efficacy of an intervention to improve CHW performance
Pineda	1983	Not a randomised controlled trial
Pradeep	2014	Non-CHW comparator
Puett	2012	Not a randomised controlled trial
Sranacharoen-pong	2012	Intervention not designed to improve performance
Turab	2014	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Zafar	2016	Intervention not designed to improve performance
Zurovac	2011	Health workers were primarily facility based

Studies awaiting classification:

Author	Year	Title
Carrón	1993	Desarrollando instrumentos de bajo costo para mejorar la calidad de los servicios rurales de planificación familiar: informe final (Enero 1992 - Diciembre 1993).
Uwima	2013	Community-based intervention to enhance provision of integrated TB-HIV and PMTCT services in South Africa.

Ongoing studies:

Hackett 2015

Trial Registration	N.S.
Trial name	Job satisfaction, self-efficacy, and performance of community health workers participating in a mobile health (mHealth) program to improve maternal, newborn and child health (MNCH) in Rural Tanzania
Objective	Evaluate the process and impact of an mHealth intervention
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: CHW pairs
Country	Tanzania
Participants	LOCAL TERM FOR CHW: Community health worker CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: mHealth KEY CHW COMPETENCY: CHWs were trained on a smartphone application designed to improve data

For peer review only

	management, patient tracking, and delivery of key health/nutrition messages to pregnant women and mothers.
Control	Ministry of Health's national community- MNCH program.
Outcomes	PATIENTS: Satisfaction CHWs: Performance
Starting date (dd/mm/yyyy)	N.S.
Contact information	Kristy Hackett PhD Candidate. University of Toronto Toronto, ON. M5T 3M7. Canada. +1-416-978-2058 kristy.hackett@mail.utoronto.ca
Funding Source	Foreign Affairs, Trade and Development Canada (DFATD) W. Garfield Weston Doctoral Fellowship Program

Källander 2012

Trial Registration	NCT01972321
Trial name	Evaluating the effect of innovative motivation and supervision approaches on community health worker performance and retention in Uganda and Mozambique: study protocol for a randomised controlled trial
Objective	Increase sustainable coverage of ICCM in Uganda and Mozambique by designing and evaluating innovations for increased CHW supervision and motivation.
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Sub-counties
Country	Mozambique and Uganda
Population	LOCAL TERM FOR CHW: Community health worker CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: mHealth KEY CHW COMPETENCY: The mHealth interventions in Uganda and Mozambique encompass three main activities: 1) closed user groups to enable free two-way communication between CHWs and their supervisors; 2) weekly ICCM data submission using phones with automated motivational feedback, SMS to supervisors flagging problems for target supervision, and summary ICCM statistics made accessible online to district statisticians; and 3) monthly motivational and constructive SMS to CHWs.
Control	Children living in the control areas continued to benefit from the routine Ministry of Health iCCM package provided by the CHWs who were supported by the national and sub-national health services with funding from Malaria Consortium's CIDA-iCCM project.

For peer review only

Outcomes	PATIENTS: Proportion of sick children appropriately treated. CHWs: (1) Performance, (2) Motivation
Starting date (dd/mm/yyyy)	11/2009
Contact information	Karin Källander, PhD +256772744126 k.kallander@malariaconsortium.org
Funding source	Bill & Melinda Gates Foundation and DFID Programme Partnership Agreement (DFID-PPA)

Rabbani 2015

Trial registration	ACTRN12613001261707
Trial name	Improving community case management of diarrhoea and pneumonia in district Badin, Pakistan through a cluster randomised study
Objective	Evaluate the affect of new supervisory strategies employed by Lady Health Supervisors
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Lady health supervisors
Country	Pakistan
Population	LOCAL TERM FOR CHW: Lady health workers CHWs: Women. Age and education level not reported. PATIENTS: Children. Sex and education level not reported.
Interventions	INTERVENTION DESCRIPTION: The intervention consists of training to build LHS knowledge and skills, clinical mentorship and written feedback to LHWs KEY CHW COMPETENCY: Community case management
Control	Standard care
Outcomes	PATIENTS: None CHWs: Improvement in CCM practices of childhood diarrhoea and pneumonia
Starting date (dd/mm/yyyy)	11/2014
Contact information	Fauziah Rabbani fauziah.rabbani@aku.edu Department of Community Health Sciences, The Aga Khan University, Stadium Road, 3500, Karachi 74800, Pakistan
Funding source	WHO Geneva, Department of Maternal, Newborn, Child and Adolescent Health.

Saha 2013

Trial registration	N.S.
Trial name (registration)	A Comprehensive Training And Supervision Program Improved Frontline Health Workers' Knowledge Of Infant And Young Child Feeding In Rural Bangladesh

For peer review only

Objective	Increase sustainable coverage of ICCM in Uganda and Mozambique by designing and evaluating innovations for increased CHW supervision and motivation.
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Not reported
Country	Bangladesh
Population	LOCAL TERM FOR CHW: Shasthya Sebika & Pushti Kormi CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: Education, Incentives, Supervision KEY CHW COMPETENCY: A comprehensive infant and young child feeding practices training and performance improvement package (3-day classroom discussion, 3-day field practice, monthly refreshers, routine supervision and performance-based incentives) plus a mass media campaign.
Control	Mass media campaign
Outcomes	PATIENTS: None CHWs: CHW knowledge of infant and young child feeding practices
Starting date (dd/mm/yyy)	2010
Contact information	Kuntal Kumar Saha, International Food Policy Research Institute (IFPRI) kuntalkumar38@gmail.com
Funding source	Bill and Melinda Gates Foundation

Vedanthan 2014

Trial registration	NCT01844596
Trial name	Optimizing linkage and retention to hypertension care in rural Kenya (LARK hypertension study): study protocol for a randomized controlled trial
Objective	Investigate whether community health workers, equipped with a tailored behavioural communication strategy and smartphone technology, can increase linkage and retention of hypertensive individuals to a hypertension care program and significantly reduce blood pressure among them.
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Community
Country	Mozambique and Uganda
Population	LOCAL TERM FOR CHW: Community health worker CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: 1) community health workers with an additional tailored behavioural

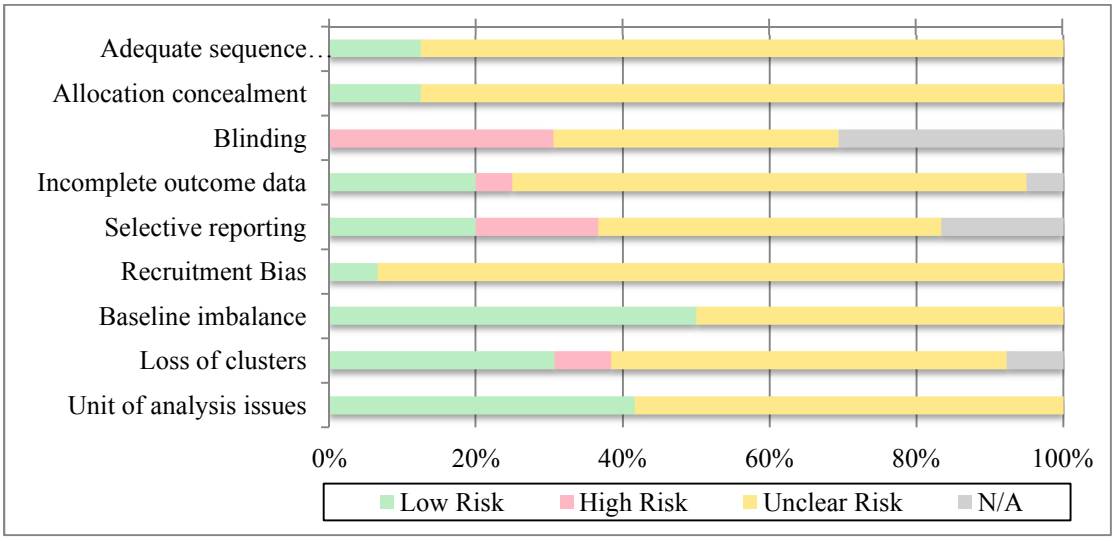
For peer review only

	communication strategy; and 2) community health workers with a tailored behavioural communication strategy who are also equipped with smartphone technology. KEY CHW COMPETENCY: Motivational interviewing and tailored communication- strategies
Control	Usual care (community health workers with the standard level of hypertension care training)
Outcomes	PATIENTS: The co-primary outcome measures are: 1) linkage to hypertension care, and 2) one-year change in systolic blood pressure among hypertensive individuals. Cost-effectiveness analysis will be conducted in terms of costs per unit decrease in blood pressure and costs per disability-adjusted life year gained. CHWs: None.
Starting date (dd/mm/yyyy)	04/2014
Contact information	Rajesh Vedanthan rajesh.vedanthan@mssm.edu Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1030, 10029 New York, USA
Funding source	National Heart, Lung, and Blood Institute of the National Institutes of Health under award number 1U01HL114200, United States, under The Global Alliance for Chronic Diseases programme

Appendix G: Risk of bias of Included Studies

Risk of bias assessments are included in the Characteristics of included studies tables in Appendix E and are summarised in the figure and table below. These assessments were used for interpreting the results and assessing the quality of evidence for specific outcomes. The first six categories apply to all included trials, while the latter four apply only to cluster-randomised control trials.

For peer review only



Review author's judgement about each methodological quality item presented as percentages across all included studies

	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Recruitment bias	Baseline imbalance	Loss of clusters	Unit of analysis issues
Andreoni (2016)	?	?	?	?	+	?				
Ashraf (2015)	?	?	?	?	-	?	+	?	?	+
Ayele (1993)	?	?	-	+	+	?	+	-	+	-
Bailey (1996)	?	?	-	?	+	?				
Bossuroy (2016)	?	?	-	+	+	?	+	+	?	+
Carrón (1994)	?	?	-	?	-	?	+	+	+	-
Chang (2011)	?	+	-	?	+	+	+	+	+	+
DeRenzi (2012A)	?	?	-	?	-	?				
DeRenzi (2012B)	?	?	-	?	-	?				
DeRenzi (2012C)	?	?	-	?	-	?				
Gautham (2015)	?	?	-	-	?	?				
Ginsburg (2014)	+	+	?	+	?	?				
Omer (1998)	+	?	-	+	+	?	+	?	+	?
Winch (2003)	?	?	-	?	?	?	+	+	+	+

peer review only

Methodological quality summary: review author's judgement about each methodological quality item for each included study

Random sequence generation and allocation concealment (selection bias)

Though all trials reported that they were randomised, only one two described the method of randomisation (Ginsburg et al., 2014; Omer et al., 2008). Likewise, only two trials reported a method of allocation concealment (Chang et al., 2011; Ginsburg et al., 2014). As such, the potential for selection bias is a notable weakness of this data set.

Blinding (performance bias and detection bias)

Given the time and human contact required for the interventions, the blinding of participants and intervention staff was not feasible. The impossibility of blinding intervention staff may have led to performance bias. Only three studies (Ayele et al., 1993; Ginsburg et al., 2014; Omer et al., 2008) reported the blinding of outcome assessors. While Chang et al. (2011) measured biological outcomes and Andreoni et al. (2016) used GPS-enabled phones meaning blinding of outcome assessors may not have been entirely crucial, the more subjective outcome collection tools used in the remaining studies (i.e. observational service tests, client questionnaires, household interviews, direct observations of CHW performance, and data collected by CHWs) make detection bias a notable risk.

Incomplete outcome data (attrition bias)

Final attrition rates for the group about which primary outcome data were gathered ranged from 0% (Ayele et al., 1993) to 30% (Derenzi et al., 2012C). Carrón et al. (1994) reported that an unspecified level of provider mortality unevenly affected the 180 CHWs randomised at baseline and attempted to compensate for this by randomly

For peer review only

deleting data from arms of the trial with less attrition until group numbers were even. Not only does this process obscure the actual attrition rate, it still leaves the study at high risk of attrition bias: randomly deleting data is not the same as never having collected it; CHWs missing from one group as a result of death may be systematically different than the ones who were randomly deleted from the other groups. Three of the studies that collected data from patients did not identify those patients a priori or include data on CHW dropout so rates of attrition are unclear (Gautham et al., 2015; Ginsburg et al., 2014; Winch et al., 2003). One of the studies, however, specified that analysis nonetheless followed the intention to treat principle, in the sense that all women in the intervention communities were analysed as though exposed, whether or not CHWs had used the intervention materials while treating them (Omer et al., 2008).

Selective reporting (reporting bias)

The variety of measurements employed in the fourteen trials suggests that no standard set of outcomes exists for the evaluation of CHW performance. Data were reported in several of formats; many were unsuitable for meta-analysis even had it been advisable. Repeated attempts were made to contact trialists for missing statistical data; while many authors were helpful in providing additional data, further information was often unavailable due to data loss or non-response. Of the trials, only two reported having a protocol (Andreoni et al., 2016; Chang et al., 2011). While Chang et al. (2011) reported all pre-specified outcomes, the outcomes listed in the protocol for the Andreoni et al. (2016) study were embargoed. As with Andreoni et al. (2016) the risk of selective reporting among the other eleven trials is thus unclear.

Publication bias

Despite an exhaustive literature search, it is possible that the findings are subject to

For peer review only

publication bias. Funnel plots were not used to explore publication bias as there were too few studies to draw conclusions from such analysis.

C-RCTs: Recruitment bias

No C-RCT recruited individuals to clusters post-randomisation, so risk of recruitment bias across trials is low.

C-RCTs: Baseline imbalance

Three studies (Ashraf et al., 2015; Carrón et al., 1994; Omer et al., 2008) did not provide an explicit statement of group equivalence at baseline, though Carrón et al. (1994) pair-matched clusters to reduce the risk of such differences. Two studies reported baseline differences: Ayele et al., (1993) reported baseline differences in CHW age and years since graduation but did not control for them; despite pair-matching clusters. Winch et al. (2003) reported baseline differences in literacy, spouse's education, and carer's education and adjusted for economic status (as measured by presence of a metal roof), education, and literacy in a multivariate random effect logistic regression model. Chang et al. (2011) and Bossuroy et al. (2016) reported that there were no significant baseline differences among participants.

C-RCT: Loss of clusters

Ashraf et al. (2015) and Bossuroy et al. (2016) did not note whether clusters were lost; in the remaining five studies, no clusters were lost.

C-RCT: Unit of analysis issues

Only three studies (Ashraf et al., 2015; Bossuroy et al., 2016; Chang et al., 2011; Winch et al., 2003) explicitly took clustering into account in the analysis, the rest were unclear or did not.

For peer review only

C-RCTs: Comparability with nonrandomised trials

Because no meta-analysis was conducted, comparability with individually randomised trials wasn't relevant.

C-RCT Contamination

Contamination and threats to programme differentiation were unlikely given the geographical separation of the different clusters and the fact that the experimental and control interventions were delivered by different personnel.

Appendix H: Cost Data

Carrón et al. (1994), Chang et al. (2011), and Derenzi et al. (2012) reported cost outcomes. Though cost was not prespecified in the review protocol as a variable about which information would be collected, it is important in assessing the relative merits of various combinations of inputs to improve CHW performance and so it is reported here Because the data on resource use and costs were very heterogeneous, meta-analysis was not appropriate and available data are presented in the table below

The economics component of this review was conducted according to current guidance on the use of economics methods in the preparation and maintenance of Cochrane reviews (Higgins & Green, 2011). The reviewer classified the included economic evaluations based on an established system (Drummond et al., 2005). Briefly, full economic evaluations (e.g. cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis) compare the costs (resource use) associated with one or more alternative interventions (e.g. intervention X versus comparator Y) with their consequences (outcomes, effects) whereas partial economic evaluations (e.g. cost analyses, cost-description studies, and cost-outcome descriptions) do not make explicit comparisons

For peer review only

between alternative interventions in terms of both costs (resource use) and consequences (effects) (Higgins & Green, 2011).

Table Summary of cost data

Primary trial (Cost paper)	Type of economic evaluation (Perspective, horizon)	Study pop.	Intervention vs. Comparator	Economic results (Currency & price year)
Carrón et al. 1994	CA (Provider, 1 year)	CHWs	I_a = Decision tree I^b = Behavioural feedback instrument $I_{a+b} = I_a + I^b$ C = Usual care	I_a was the least expensive (29.42 USD/CHW) followed by I^b (78.45 USD/CHW), the I_{a+b} (106.80 USD/CHW), and C, the traditional training course (166.17 USD/CHW in groups of twenty-five, 118.40 USD/CHW in groups of fifteen or fewer) (1993 USD)
Chang et al. 2011	CA (Provider, 2 years)	HIV positive adults	I = A text message reporting system and health centre warmline C = Usual care	The total cost of the mHealth intervention was $2,353.24. The yearly cost was $1045.88, resulting in a yearly additional cost per patient of $2.35. (2006-08 US)
DeRenzi et al., 2012B	CA (Provider, 1 year)	Chronic care patients	I = Escalating text message reminder system & follow-up by supervisor C = Usual care	The intervention adds an estimated $0.84 per client per year of running the ICT4CHW programme with automated escalating reminders (Year unspecified USD)

CA: Cost analysis; **USD**: US dollars.

These studies generally highlight the low incremental costs or potential cost savings of these CHW performance interventions. The quality of the economic evaluations was assessed using the Cochrane-recommended CHEC list (Evers et al., 2005), a consensus criteria list that incorporates items from Drummond and Jefferson (1996) and other systematically-identified guidelines. These assessments are reported in the table below. Because CHEC was designed to assess full economic evaluations, items not applicable to the partial economic evaluations included are shaded grey. There was some risk of bias across the studies: not all relevant alternatives and costs (e.g. training, supervision) were reported in all studies, studies did not perform discounting, and no studies adequately addressed ethical and distributional implications.

Table Quality of included economic evaluations

	Item	Carrón 1994	Chang 2011	DeRenzi 2012ABC
1	Is the study population clearly described?	+	+	+
2	Are competing alternatives clearly described?	+	+	+
3	Is a well-defined research question posed in answerable form?	-	-	-
4	Is the economic study design appropriate to the stated objective?			
5	Is the chosen time horizon appropriate to include relevant costs and consequences?			
6	Is the actual perspective chosen appropriate?	+	+	+
7	Are all important and relevant costs for each alternative identified?	+	+	-

For peer review only

8	Are all costs measured appropriately in physical units?	+	+	+
9	Are costs valued appropriately?	+	+	+
10	Are all important and relevant outcomes for each alternative identified?			
11	Are all outcomes measured appropriately?			
12	Are outcomes valued appropriately?			
13	Is an incremental analysis of costs and outcomes of alternatives performed?			
14	Are all future costs and outcomes discounted appropriately?	-	-	-
15	Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	-	-	-
16	Do the conclusions follow from the data reported?	+	+	+
17	Does the study discuss the generalizability of the results to other settings and patient/ client groups?			
18	Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?	+	+	-
19	Are ethical and distributional issues discussed appropriately?	-	-	-

References for Appendix H

Higgins, J. P. T., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*: The Cochrane Collaboration.

Drummond, M., & Jefferson, T. (1996). Guidelines for authors and peer reviewers of economic submissions to the BMJ. *British Medical Journal, 313*(7052), 275-283.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford, UK: Oxford University Press.

Evers, S., Goossens, M., de Vet, H., van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *Int J Technol Assess Health Care, 21*(2), 240-245.

For peer review only

Appendix I: PRISMA Checklist

SECTION/TOPIC	#	CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	2
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Appendix A
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4-5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-6 and Appendix C
Search	8	Present full electronic search strategy for at least one database, including any	6 and Appendix C

For peer review only

		limits used, such that it could be repeated.	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Appendix D
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6-7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	6-7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7-8
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period)	9-12 and Appendix E

For peer review only

		and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Appendix E & G
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	12-19
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	12-19
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	20
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	24
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	21-25
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	25

For peer review only

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	2
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Appendix A
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4-5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-6 and Appendix C
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6 and Appendix C
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Appendix D
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6-7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

PRISMA 2009 Checklist

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	6-7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7-8
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	9-12 and Appendix E
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Appendix E & G
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	12-19
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	12-19
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	20
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	24
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	21-25
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	25

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

PRISMA 2009 Checklist

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Page 2 of 2

For peer review only

BMJ Open

A Systematic Review of Interventions for Improving the Performance of Community Health Workers in Low- and Middle-Income Countries

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-014216.R1
Article Type:	Research
Date Submitted by the Author:	20-Jun-2017
Complete List of Authors:	Ballard, Madeleine; University of Oxford, Department of Social Policy and Intervention Montgomery, Paul; University of Birmingham, Department of Social Policy and Social Work
Primary Subject Heading:	Global health
Secondary Subject Heading:	Evidence based practice, Health policy, Public health
Keywords:	Human resource management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, PUBLIC HEALTH, Community child health < PAEDIATRICS, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts

A Systematic Review of Interventions for Improving the Performance of Community Health Workers in Low- and Middle-Income Countries

Madeleine Ballard, BA, MSc, DPhil[a]
Paul Montgomery, BA, MSc, DipSw, DPhil[b]

[a] Centre for Evidence-Based Intervention, University of Oxford
32 Wellington Square, Oxford OX1 2ER, UK
[b] Department of Social Policy and Social Work, University of Birmingham
Muirhead Tower, Ring Road, Birmingham B15 2TT

For peer review only

Correspondence:

Madeleine Ballard
Centre for Evidence-Based Intervention
University of Oxford
32 Wellington Square
Oxford OX1 2ER, UK
E: madeleine.ballard@gmail.com
P: +49 15255604201

ABSTRACT

Objective: To systematically review and critically appraise the evidence for the effects of interventions to improve the performance of community health workers (CHWs) for community-based primary healthcare in low- and middle-income countries.

Design: Systematic review following PRISMA guidelines.

Methods: Nineteen electronic databases were searched with a highly sensitive pre-specified strategy and the grey literature examined, completed July 2016. Randomised controlled trials evaluating interventions to improve CHW performance in LMICs were included and appraised for risk of bias. Outcomes were biological and behavioural patient outcomes (primary), utilisation of health services, quality of care provided by CHWs, and CHW retention (secondary).

Results: Two reviewers screened 8082 records; fourteen evaluations were included. Due to heterogeneity and lack of clear outcome data, no meta-analysis was conducted. Results were presented in a narrative summary. The review found one study showing no effect on the biological outcomes of interest, though these moderate quality data may not be indicative of all biological outcomes. It also found moderate quality evidence of the efficacy of performance improvement interventions for (i) improving behavioural outcomes for patients, (ii) improving utilisation of services by increasing the absolute number of patients who access services and, perhaps, better identifying those who would benefit from such services, and (iii) improving CHW quality of care in terms of upstream measures like completion of prescribed activities and downstream measures like adherence to treatment protocols. Nearly half of studies were compound interventions, making it difficult to isolate the effects of individual performance improvement intervention components, though four specific strategies pertaining to recruitment, supervision, incentivisation, and equipment were identified.

Conclusions: Variations in recruitment, supervision, incentivisation, and equipment may improve CHW performance. Practitioners should, however, assess the relevance and feasibility of these strategies in their health setting prior to implementation. Component selection experiments on a greater range of interventions to improve performance ought to be conducted.

Keywords: Community health workers; lay health workers; supervision; incentives; recruitment; training; systematic review

Strengths and limitations of this study

- This is the first systematic review of interventions for improving the performance of community health workers that conducts risk of bias assessment, assessment of the quality of the body of evidence, and follows the PRISMA reporting guidelines.

For peer review only

- The review strengths include its pre-specified, systematic, and highly sensitive search strategy; inclusion of published and unpublished literature; and inclusion of only the most rigorous evidence.
- Though the review has carefully specified the inclusion criteria in keeping with previous studies, there is no widely accepted definition for this cadre of health worker and so some exclusions may be debated.
- Poor reporting and nonresponse from contacted authors meant it was not possible to obtain some relevant missing data on methodological characteristics, clinical characteristics, and outcomes for some trials
- Given mounting pressure to meet ambitious new international health goals and avoid repeating the mistakes of large-scale CHW programs of the past, the review is timely.

For peer review only

INTRODUCTION

Community health workers (CHWs)—lay workers to whom simple medical procedures can be "task shifted" from higher-level medical providers (e.g. nurses, doctors)—were widely promoted as a means to provide primary healthcare in resource poor settings as early as the 1978 Alma-Ata Declaration [1]. A series of reviews in the late 1980s and early 1990s, however, found that large-scale CHW programs often failed to replicate the success of smaller community-based programs [2-7].

Rigorous evidence has since accumulated on the efficacy of CHWs to deliver assorted health interventions [8-11] and—prompted by the failure to meet the health-related Millennium Development Goals [12, 13] and a continuing global health worker shortage that imperils the new Sustainable Development Goals (SDGs) [14]—CHWs have once again been proposed as a way to extend services to hard-to-reach populations in remote areas [15-19].

The most recent evaluations of national-scale CHW programs, however, remain unfavourable [20-22]. Yet, it is difficult to assess on the basis of such evaluations whether a given CHW program did not achieve a statistically significant effect because (i) such programs do not work, (ii) the CHW program was not

implemented properly (type III error), or (iii) the CHW program design was not yet optimised to achieve maximum effect (e.g. via the best or most efficient combination and dose of intervention components) [23-26]. Given that qualitative assessments suggest that poor outcomes are associated with sub-optimal program design [2, 27-30], a systematic review of studies disaggregating the effect of individual intervention components could help improve program design in future [24].

While history has demonstrated that broad contextual factors related to the health system or the political and economic climate can have an impact on CHW performance [31], this review is focused on how modifications to aspects of the delivery mechanism itself—namely, CHW program components (e.g. recruitment, supervision)—can be used to optimise CHW performance.

METHODS

Inclusion criteria

Randomisation allows the effects of particular program components to be isolated; this is a critical step in intervention optimisation and the goal of this review (see protocol in Appendix A) [24]. Because health workforce outcomes are empirically testable and workplace interventions possible to randomise, eligible trials were randomised controlled trials, cluster-randomised trials, crossover trials, and factorial trials. Studies were only included where the primary objective, or one of the primary objectives, was to determine the efficacy or effectiveness of an intervention to improve the performance of CHWs. This includes 'head-to-head' comparison studies of CHWs engaging in the same task (e.g. providing ART adherence support) in the context of different program designs (e.g. more or less supervision) that have to date been excluded from reviews of CHW effectiveness [e.g. 10].

Several definitions for and variants of the term CHW have been employed in

For peer review only

the literature. For the purpose of this review, CHWs were defined as: any lay health workers who (a) live in the catchment they serve, (b) are primarily based in the community (as opposed to a health facility), (c) belong to the formal health system (i.e. are managed by the government or an implementing NGO), (d) perform tasks related to healthcare delivery, and (e) have received organised training but have no formal or paraprofessional certification or tertiary education degree [10, 32]. Aside from residence in a low- or middle-income country (LMIC) as classified by The World Bank Group (listed in Appendix B), there were no restrictions on the types of patients for whom data were extracted.

Any intervention designed to improve CHW performance, compared to CHWs who did not experience the intervention, was included. Studies that were designed to examine the efficacy of a particular therapy (e.g. zinc tablets for diarrhoea), as opposed to the relative effectiveness of different ways of designing programs that support CHWs to deliver said therapy were excluded. Interventions of any duration or follow up were included.

A number of processes and outcomes might be affected by interventions that aim to improve the performance of CHWs. Outcomes were extracted and categorised as biological and behavioural patient outcomes (primary), utilisation of health services, quality of care provided by CHWs, and CHW retention (secondary). Outcomes of interest and study quality were not used as criterion for screening studies.

Search Strategy

A comprehensive search strategy was used to identify all relevant studies regardless of language, year, or publication status (published, unpublished, in press, and in progress). We searched:

For peer review only

1. 3ie Impact Evaluation Database
2. AMED
3. BiblioMap
4. British Library for Development Studies at IDS
5. CINAHL
6. CENTRAL
7. DoPHER
8. EMBASE
9. Global Health
10. IDEAS
11. IPA Database
12. Databases within ISI Web of Science
13. JOLIS
14. J-PAL Evaluations Database
15. MEDLINE
16. MEDLINE In-Process & Other Non-Indexed Citations
17. POPLINE
18. Pubmed
19. World Bank IEG
20. WHOLIS

The following strategy was refined in consultation with an information specialist and used to search MEDLINE. In order to maximise sensitivity, no randomisation filter was included [33]. (Precise dates of coverage, date last searched, search strategies used for the databases listed above, and details of other resources searched, including the grey literature, can be found in Appendix C.)

1. ((community or village) adj2 (agent$ or aid$ or promot$ or mobili?er$ or distribut?r$ or worker$)).ti,ot,ab,kf.
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliary$ or extension) adj2 (worker$ or volunteer$)).ti,ot,ab,kf.
3. (accompanier$ or accompagnateur$ or activista$ OR animatrice$ OR brigadista$ or kader$ or monitora$ or promotora$ or sevika$ or fhw$ or chw$ or lhw$ or vhw$ OR chv$ or "shastho shebika" or "shasto karmis" or anganwadi$ or "barefoot doctor" OR "agente comunitario de salud" or "agente communitario de saude").ti,ot,ab,kf.
4. (performance or effective$ or skill$).ti,ot,ab,kf.
5. 1 or 2 or 3
6. 5 and 6

Searches were conducted May 4, 2014 and updated ending July 12, 2016.

Data collection and analysis

Two authors independently screened titles, abstracts, and full text articles. Included articles were determined by consensus. The extraction sheet was informed by the Cochrane EPOC Group data collection checklist [34], the Template for

For peer review only

Intervention Description and Replication [35], and the Oxford Implementation Index [36] (see Appendix D). Reviewers extracted data on study design, participants, interventions, outcomes and methodological quality.

Risk of bias for all RCTs was assessed using The Cochrane Collaboration's tool for evaluating the risk of bias [37]. The quality of evidence across a body of evidence (i.e., multiple studies with similar interventions and outcomes) was assessed using the GRADE approach [38]. These assessments are made with the caveat that—as had been noted elsewhere [39-41]—there are issues with the applicability, reproducibility, and clarity of GRADE in the field of public health and with complex interventions.

Cluster-randomised trials were verified to ensure appropriate analysis had been done (i.e. adjustment for clustering); if such an analysis had not been done, the necessary intracluster correlation coefficients (ICCs) were extracted or obtained from the investigators and results reanalysed [39]. All comparisons from factorial studies were included.

RESULTS

Characteristics of included studies

Excluding duplicates, a total of 8082 records were screened for inclusion (Fig 1). A total of 12 records reporting 14 studies were included [42-53]. Complete information for each trial is provided Appendix E; study characteristics are described in Table 1 [35]. Information on excluded, ongoing and studies awaiting classification is provided in Appendix F.

[INSERT IMAGE 1]

Figure 1 PRISMA study flow diagram

For peer review only

Ten countries were represented: Bangladesh, Ethiopia, Guatemala, India (2), Mali, Pakistan (2), Paraguay, Uganda, Tanzania (3), Zambia. Three trials evaluated supervision interventions [43ABC], two incentive interventions [46, 53], two equipment interventions [42, 44], one a recruitment intervention [52], one a training intervention [45], and five "compound" interventions altering more than one program component [47-51]. No information about implementation fidelity—i.e. whether an intervention is delivered as intended by the developers—was provided [23, 36]. The reporting of CHW and patient characteristics and contextual factors was poor. Assessments of the risk of bias for included studies are included in Appendix G and are considered when assessing the findings.

For peer review only

Table 1 Brief Study Characteristics

Study	Setting (Study Design)	CHW & Patient Sample	Intervention Type & Description of Trial Arms	Outcomes (Follow-up)
Andreoni 2016 [46] **Objective:** Test the effect of tailored contracts for CHWs conducting door-to-door immunisation drives in rural and urban Pakistan	Rural & Urban Pakistan (2x2 factorial RCT)	Female CHWs (I_{a+c}=85, I_{a+d}=84, C_{b+c}=88, C_{b+d}=80). Patients: children under 5 yrs. (N= unspecified)	**Incentives** I: a+c; a+d C: b+c; b+d (a) Employment contracts tailored to individual discounting preferences (b) Untailored contracts (c) Advanced allocation of daily work targets (d) Day-of allocation	**CHWs:** Closeness to vaccine distribution policy objective **Patients:** None (Immediate)
Ashraf 2015 [52] **Objective:** Test the effect of advertising career possibilities at the recruitment stage for CHWs providing basic primary care	Rural Zambia (C-RCT)	CHWs (I=149, C=149). Patients: children under 5 yrs. & women (N = unspecified)	**Recruitment** I: Recruitment materials emphasising career possibilities C: Recruitment materials emphasising benefits to community	**CHWs:** Home visits, patients seen, comm. mtgs organised **Patients:** Use of services (18 months)
Ayele 1993 [51] **Objective:** Test the effect of an in-person refresher course and	Rural Ethiopia (C-RCT)	Male CHWs (I=50, C=50). Patient characteristics not reported (N= unspecified)	**Training-Supervision** I: 5 day in-person refresher course and monthly supervision C: No refresher, no supervision	**CHWs:** Composite functional status score **Patients:** None (6 months)

monthly "supportive" supervision on CHW-led health promotion				
Bailey 1996 [45] **Objective:** Test the effect of tiered, interactive training featuring visual cue cards in improving CHW case management of diarrhoea	Rural Guatemala (RCT)	CHWs (I=22, C=27). Patient characteristics not reported (N= unspecified)	**Training** I: Interactive, tiered training with visual cue cards C: Didactic training	**CHWs:** Knowledge score **Patients:** None (Immediate)
Bossuroy 2016 [53] **Objective:** Test the effect of performance-based incentives on CHWs providing TB case detection and adherence counselling	Urban India (C-RCT, 4 arm)	CHWs (N=78). Adult patients (N = 2500).	**Incentives** I_a: Financial incentives based on patient detection for six months then based on treatment adherence I_b: Financial incentives based on patient detection for six months then a fixed salary I_c: Fixed salary for six months then incentives based on treatment adherence C: Fixed salary	**CHWs:** TB case detection **Patients:** TB default rate (6 months)
Carrón 1994 [50] **Objective:** Test the effect of an interactive counselling guide and supervisory behavioural	Rural Paraguay (2x2 factorial C-RCT)	Female CHWs (I_a=45, I_b=45, I_{a+b}=45, C=45). Patient characteristics not reported (N=unspecified)	**Training-Supervision** I: a; b; a+b (a) Trained using interactive service algorithm (Decision tree) (b) Role-play based supervision	**CHWs:** Quality of care provided **Patients:** None (12 months)

feedback instrument on CHWs distributing family planning materials			(Behavioural feedback instrument) C: Usual training and supervision	
Chang 2011 [48] **Objective:** Test the effect of a mobile phone CHW support intervention on HIV care	Rural Uganda (C-RCT)	CHWs (I=13, C=16). Adult HIV+ patients (I=446, C=524)	**Integration-Equipment** I: A text message reporting system and health centre hotline C: No mobile health intervention (usual care)	**CHWs:** None **Patients:** Virologic failure, adherence (2 years)
DeRenzi 2012A [43] **Objective:** Test the effect of SMS reminders on the promptness of routine CHW visits	Urban Tanzania (Pilot RCT)	Female CHWs (I=8, C=7). Patient characteristics not reported (N= unspecified)	**Supervision** I: Escalating SMS reminder system C: No reminders	**CHWs:** None **Patients:** Change in closed referrals (24 days)
DeRenzi 2012B [43] **Objective:** Evaluate the impact of SMS reminders plus escalation to supervisor on the promptness of routine CHW visits	Urban Tanzania (RCT)	CHWs (I=34, C=40) Chronic care patients (N=unspecified)	**Supervision** I: Escalating SMS reminder system & supervisor follow-up for underperforming CHWs C: No reminders	**CHWs:** None **Patients:** Mean time overdue at clinic (40 days)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

DeRenzi 2012C [43] **Objective:** Test the effect of escalating to the supervisor versus SMS messages alone on the promptness of routine CHW visits	Urban Tanzania (RCT)	CHWs (N=87) Chronic care patients (N=unspecified)	**Supervision** I: Escalating SMS reminder system & supervisor follow-up for underperforming CHWs C: Escalating SMS reminder system only	**CHWs:** None Patients: Mean time overdue at clinic (3 months)
Gautham 2015 [42] **Objective:** Test the effect of mobile phone-based, media-rich procedural guidance applications on CHWs providing primary care	Rural India (RCT)	CHWs (I=8, C=8). Paediatric and adult patients (I=65, C=61)	**Equipment** I: Mobile phone–based, media-rich procedural guidance applications C: Paper guidelines	**CHWs:** Protocol compliance **Patients:** None (2 months)
Ginsburg 2014 [44] **Objective:** Test the effect of mobile guidance applications on CHWs referring women with abnormal breast exams	Rural Bangla-desh (3-arm RCT)	Female CHWs (I_a=10, I_b=10, C=10). Female adult patients (I_a=246, I_b=240, C=70)	**Equipment (I_a)/** **Training-Equipment (I_b)** I: a; b (a) Smart phone with applications to guide interview, report data, show motivational video, and offer an appointment (b) As in (a) plus additional "patient navigation training." C: Paper-based tools	**CHWs:** None **Patients:** Adherence with referral (1 month)
Omer 2008 [47] **Objective:** Test the	Urban Pakistan (C-RCT)	Female CHWs (I=52, C=unspecified). Mothers and their children (I=529,	**Integration-Training** I: Indigenous, evidence-based communication tool	**CHWs:** Home visits **Patients:** Uptake of protective health

For peer review only

effect of an indigenous, evidence-based communication tool on CHWs providing maternal care		C=541)	C: Usual care	practices (10 months)
Winch 2003 [49] **Objective:** Test the effect of visual aids, refresher training, a counter referral recordkeeping notebook, and a community meeting on CHWs treating common illnesses	Rural and Urban Mali (C-RCT)	CHWs (N=~102). Children <14 yrs. (N=152, C=134)	**Integration-Training-Reporting** I: Visual aids, refresher training, a counter referral recordkeeping notebook, a community meeting C: Standard training	**CHWs:** Proportion of children treated at home **Patients:** Closed referrals

$\bar{x}$**:** mean; **C**: control; **Comm:** community; **C-RCT:** cluster RCT; **I:** intervention; **N:** intervention/control participant breakdown unspecified, total number of participants; **mtg:** meeting; **RCT:** randomised controlled trial; **SMS**: short messaging service (i.e. text message); **yrs**: years

For peer review only

Effects of Interventions

Results are presented in a structured summary [36]. Where possible, standardised measures of effect (RR, SMD) have been provided to aid comparison. Due to missing data, multiple outcome measures were not possible to standardise and are reported as in the original trial papers. Although six authors were contacted regarding missing data, none were able to supply it. The protocol for this review (see Appendix A) specified that a meta-analysis would be conducted if appropriate. Meta-analysis would, however, be uninformative due to heterogeneity and lack of clear outcome data.

Five studies examined the review's primary outcomes [43BC, 47-49] and nine studies examined secondary outcomes [42, 43A, 44-47, 50-52].

Primary outcomes

The primary outcomes specified in the protocol were objective physical outcomes for patients and subjective health outcomes for patients (Table 2). Overall, these studies suggest that CHW performance improvement interventions can improve certain behavioural, but not biological, outcomes for patients. The body of evidence for behavioural outcomes for patients was graded moderate (downgraded for design and implementation as allocation concealment and random sequence generation were unclear in the majority of studies) as was the evidence of no effect for biological outcomes (downgraded for design and consistency as the result is based on only one study that may be underpowered). The data for both outcomes is considered in detail below.

For peer review only

Table 2 Primary outcomes

Study (N)	Intervention Type	Variables	Measure of Effect (95% CI or SE)
Biological outcomes for patients			
Chang 2011 (N=970 patients)	Reporting-Supervision	Virologic failure	RR = 1.17 (0.84–1.64) P = 0.34
		Mortality	RR = 0.82 (0.55–1.22) P = 0.33
Behavioural outcomes for patients			
Bossuroy 2016 (N = 2500 patients)	Incentives	TB default rate	B = -0.04 (SE = 0.06) P = 0.25
Chang 2011 (N=970 patients)	Reporting-Supervision	<95% pill adherence	RR = 1.01 (0.97–1.06) P = 0.59
		Loss to follow-up	RR = 1.29 (0.50–13.32) P = 0.60
DeRenzi 2012B (N=un-specified patients)	Supervision	Mean days clients overdue	U = 271.00 P < 0.001 r = 0.500
DeRenzi 2012C (N=un-specified patients)	Supervision	Mean days clients overdue	See text
Omer 2008 (N=1070 patients)	Integration-Training	Attend at least one prenatal visit	RR = 1.94 (1.56-2.41) P < 0.0001
		Colostrum to newborns	RR = 1.21 (1.12-1.32) P < 0.0001
		Stop routine heavy work	See text
		Exclusive breastfeeding	See text
Winch 2003 (N=286 patients)	Training-Reporting-Integration	Proportion of malarial children treated by mothers exactly per IMCI norms	RR = 27.76 (0.53-1441.23) P < 0.0001

B: regression coefficient; **C:** control; **I:** intervention; **P:** p-value; **r:** regression coefficient; **RR**: risk ratio; **SE:** standard error; **U:** Mann-Whitney U-statistic

Only one study reported biological outcomes for patients [48], finding no significant differences in HIV patients' cumulative risk of virologic failure (RR = 1.17, 95% CI 0.84–1.64, P = 0.34) or mortality (RR = 0.82, 95% CI 0.55–1.22, P = 0.33).

Six studies reported behavioural outcomes for patients: Bossuroy [53] found no significant difference in patient TB default rates (B = -0.04, SE = 0.06, P = 0.25), though CHWs were incentivised based on treatment adherence. Chang [48] found no significant differences in HIV patients'<95% pill count adherence (RR = 1.01, 95% CI 0.97–1.06, P = 0.59) or loss to follow-up (RR = 1.29, 95% CI 0.50–13.32, P = 0.60). Omer [47] found that pregnant women in intervention communities were more likely to attend at least one prenatal check-up (RR = 1.94, 95% CI 1.56-2.41, P < 0.0001), give colostrum to their newborn babies (RR = 1.21, 95% CI 1.12-1.32, P < 0.0001), and maintain exclusive breastfeeding for four months but equally likely to stop routine heavy work.[1] No quantitative data was provided for the latter two outcomes. Winch [49] found that the proportion of malarial children treated exactly per IMCI norms was significantly higher in the intervention group than the control group (RR = 27.76, 95% CI 0.53-1441.23, P < 0.0001). DeRenzi [45B] found that the mean number of days patients were overdue at the clinic was lower in the intervention than the control group (U = 271.00, P <0.001, r = 0.50). In DeRenzi [45C], control CHWs who had been in the SMS+Supervision group during DeRenzi [45B] performed worse relative to those who continued to receive SMS+Supervision (U = 68.00, P = 0.023, r = 0.340). The performance of CHWs who had been in the control group during DeRenzi [45B] improved; however, no statistically significant difference was found between the SMS+Supervision and SMS-only groups (U =

[1] A woman advised by a CHW to reduce routine heavy work, however, was more likely to adhere to the advice than someone who was not so advised or who was advised by others (RR = 1.31, 95% CI 1.20 – 1.422, P < 0.0001).

For peer review only

101.00, P = 0.880, r = 0.043).

Secondary outcomes

Secondary outcomes were utilisation of health services, quality of care provided by CHWs, and CHW retention (Table 3). Overall, these studies suggest that CHW performance interventions can improve utilisation of services by increasing the absolute number of patients who attend the health centre and, possibly, by better identifying those who would benefit from such services. They also suggest that such interventions can improve CHW quality of care in terms of upstream measures like completion of prescribed activities and downstream measures like adherence to treatment protocols and observed differences in the calibre of care provided. The body of evidence for both outcomes was graded moderate (utilisation downgraded for design and implementation as allocation concealment, random sequence generation, and blinding of assessors as unclear in the majority of studies; quality of care downgraded for design and implementation as allocation concealment, random sequence generation, and blinding of assessors were unclear in the majority of studies).

No studies reported on CHW retention or adverse events. Though cost was not prespecified in the review protocol as a variable about which information would be collected, it is important in assessing the relative merits of various combinations of inputs to improve CHW performance and so it is reported in Appendix H. The data for utilisation and quality of care is considered in detail below.

Table 3 Secondary outcomes

For peer review only

Study	Intervention Type	Variables	Measure of Effect (95% CI or SE)
Utilisation of health services			
Ashraf 2015 (N = specified in text)	Recruitment	Utilisation of health services	See text
DeRenzi 2012A (N=11 CHWs)	Supervision	Closed referrals per CHW	SMD = 0.86 (-0.30, 2.03)
Ginsburg 2014 (N=556 patients)	Equipment (I_a) /Training-Equipment (I_b)	Closed referrals	I_a: 43% I_b: 63% C: 53% P > 0.05
Winch 2003 (N=79 patients)	Training-Reporting-Integration	Closed referrals	RR = 0.78 (0.57, 1.07) P = 0.13
CHW quality of care			
Andreoni 2016 (N = 337 CHWs)	Incentives	Closeness to vaccine distribution policy objective	SMD = - 0.32 (-0.54, -0.1)
Ashraf 2015 (N= 298)	Recruitment	Home visits	B = 93.95 (SE = 37.19) P = 0.006
		Patients seen	B = 31.79 (SE = 260.4) P = 0.45
		Comm. mtgs organised	B = 17.06 (SE = 5.22) P < 0.001
Ayele 1993 (N=100 CHWs)	Training-Supervision	Completion of government-set CHW activities	See text
Bailey 1996 (N = 49)	Training	Knowledge score	Combined I: 36.8% C: 24.9% P > 0.05
Bossuroy 2016 (N = 78 CHWs)	Incentives	TB case detection rate	See text
Carrón 1994 (N=180 CHWs)	Training-Supervision	Quality of care score	See text
Gautham 2015	Equipment	Protocol compliance	See text

For peer review only

(N= 16 CHW)			
Winch 2003 (N=102 CHWs)	Training-Reporting-Integration	Appropriate drug administration	RR = 1.45 (0.99, 2.13) P = 0.04
		Dispensing correct amount of medication	RR = 1.22 (0.69, 2.18) P = 0.47

B: regression coefficient; **C:** control; **I:** intervention; **P:** p-value; **RR**: risk ratio; **SE:** standard error; **SMD:** standardized mean difference

Four studies reported on utilisation; three examined the effect of CHW performance interventions on closed referrals [43A, 44, 49] and one on utilisation more generally [52]. Winch [49] found no difference in the rate of compliance with referral to the health centre among intervention or control patients (RR = 0.78, 95% 0.57 to 1.07, P = 0.13) but that patients in the intervention group were nearly four times more likely to be referred than those in the control group (RR=3.76, 95% CI 2.25 to 6.28, P <0.0001). Ginsburg [44] found similar results: while women in intervention arm B (smart phones plus patient navigation) were significantly more likely to attend for care versus women in intervention arm A (smart phones without navigation; 63% vs. 43%, P < 0.0001), adherence in the two intervention arms combined was the same (53%) as that of the control arm. Control arm CHWs interviewed fewer women than CHWs in either of the two intervention arms, however, and the proportion of study participants found to have an abnormal clinical breast exam was more than 3 times greater in the intervention arms (I_A=3.1%, I_B =3.2%) arms than the control arm (1.0%) (P < 0.0001). DeRenzi [45A], by contrast, found a large but statistically insignificant increase in the mean percentage of closed referrals in the intervention group (SMD = 0.86, 95% CI -0.30, 2.03), though a methodological limitation of the trial may explain much of the difference. Ashraf [52] found the intervention increased the number of women giving birth at the health centre by 31% (N = 1269 patients, B = 14.68, SE = 6.32, P = 0.01), the number of

For peer review only

children under 5 undergoing health checks by 24% (N = 1618 patients, B = 318.1, SE = 98.05, P < 0.001), the number of children under 5 being weighed by 23% (N = 1610 patients, B = 284.3, SE = 110.2, P = 0.005) and number of children under 1 receiving immunization against polio by 20% (N = 1530 patients, B = 14.98, SE = 4.803, P < 0.001). No statistically significant difference was seen in the number of postnatal (0-6 week) visits, children under 1 year receiving BCG vaccinations, and children under 1 receiving measles vaccinations.

Eight studies reported on CHW quality of care using a range of measures [42, 45, 46, 49-53]. Four trials reported on CHWs completing [51-53] or having the knowledge necessary to complete [45] prescribed activities. Ayele [51] calculated the scores for the 13 CHW government-set activities (e.g. home visits, referrals), finding that 10 of 13 activity scores (similar at baseline, maximum score not stated, P < 0.05) and composite functional status score (P < 0.0001) were significantly higher among intervention group CHWs. Though the composite functional status score was defined as the sum of the 13 activity scores, this is not the case. The authors of the paper were contacted but were unable to supply clarifying data. Ashraf [52] found that CHWs recruited with career possibilities salient conduct 29% more household visits (B = 93.95, SE = 37.19, P = 0.006) and organise over twice as many community meetings (43 vs. 22, B = 17.06, SE = 5.220, P < 0.001) than those recruited with community benefits salient, while the difference in the number of patients seen at the health post is also positive but not significant (B = 31.79, SE = 260.4, P = 0.45). Bossuroy [53] Found that, on average, the number of new TB detections increased by 2.18 (33.2%) each month among intervention CHWs during the period they were incentivised based on case detection (B = 2.18, SE = 0.95, P = 0.01). The number of defaults, however, was significantly larger among intervention CHWs during this period (B = 0.08, SE =

For peer review only

0.04, P = 0.02). Bailey [45] found that CHWs in the intervention group demonstrated a significantly greater ability to correctly diagnose diarrhoea of varying type and severity (I = 77.3%, C = 43.1%, P < 0.05). Though non-significant differences favouring the intervention group were found in CHWs' diarrhoea knowledge, referral recommendations, and treatment practices.

Two studies reported on treatment protocol compliance [42, 49]. Winch [49] found that intervention CHWs were significantly more likely to appropriately administer drugs (RR = 1.45, 95% CI 0.99 to 2.13, P = 0.04). The study also reported that more intervention CHWs (58.9%) sold the correct amount of medication than control CHWs (48.1%). Using the ICC provided (0.210), a risk ratio was approximated (RR = 1.22, 95% CI = 0.69 to 2.18, P = 0.47). The p-value calculated by the authors (P = 0.042) indicates statistical significance; emails attempting to resolve the discrepancy received no response. Gautham [42] calculated mean protocol adherence for the intervention and control groups across a variety of stratifications of interest. The intervention group exhibited greater protocol compliance across patient order and across sexes than the control group, though the initial protocol compliance superiority of the intervention group diminished over time. Authors were contacted for the unstratified mean protocol compliance scores but were not able to provide them.

Two trials reported on observed differences in care provided [46, 50]. Carrón [50] found quality of care provided improved in all four groups (P < 0.05). The largest quality gain was observed in CHWs who received the decision tree, though differences between groups were not statistically significant. Measures of variability were not reported. Andreoni [46] found that, when outliers were excluded, CHWs who receive tailored contracts were one-third closer to the policy objective than their

For peer review only

untailored counterparts (SMD = -0.32, 95% CI -0.54 to -0.1) with an even greater effect for day-of allocation (B = 0.130, SE = 0.037, P < 0.001).

DISCUSSION

Summary of Main Results

This review identified fourteen trials evaluating the effects of interventions for improving the performance of CHWs for community-based primary healthcare in LMICs. A structured summary laid out moderate quality evidence of the efficacy of CHW performance interventions in (i) improving certain behavioural outcomes for patients, (ii) improving utilisation of services by increasing the absolute number of patients who attend the health centre and, possibly, by better identifying those who would benefit from such services, and (iii) improving CHW quality of care in terms of upstream measures like completion of prescribed activities and downstream measures like adherence to treatment protocols and observed differences in the calibre of care provided (e.g. technical competence). The evidence for biological outcomes was assessed as moderate quality, though the lack of effect on virologic failure and mortality may not be indicative of all possible biological outcomes. None of the studies included in this review reported on possible adverse effects of these interventions, either to patients, CHWs, or the health system. The heterogeneity and poor reporting of included studies precluded meta-analysis. Evidence is insufficient to draw conclusions regarding the effects of such interventions on retention or to specify which performance improvement intervention strategies are likely to be most effective. Furthermore, the nearly half of studies were compound interventions, making it difficult to isolate the effects of individual performance improvement intervention components.

For peer review only

That said, there is moderate quality evidence that the following practices improve behavioural outcomes for patients, utilisation of services, and/or CHW quality of care: (1) When recruiting CHWs, emphasising career possibilities rather than benefits to the community [52]. (2) When supervising CHWs, providing escalating reminders for tasks that are overdue and following-up with underperforming CHWs [43ABC]. (3) When incentivising CHWs, tailoring incentives to measure individual preferences—but only for CHWs performing a single repetitive task, not for CHWs who must perform multiple or more complex tasks [46, 53]. (4) When equipping CHWs, using mobile phone-based procedural guidance applications [42, 44]. These four strategies ought to be interpreted with caution, however, for reasons discussed in the following two sections.

Internal validity and implications for research

Methodological quality of the fourteen included studies was difficult to assess due to incomplete reporting of key methodological and clinical features; contacting the trialists did not always yield sufficient additional information to make such judgments. Missing data and unit of analysis problems made the size [47] and precision [49-51] of several outcomes difficult to ascertain.

The selection and validity of outcome measures is also a concern across studies. For example, several measurement instruments were employed to assess "quality of care." Some were more valid than others: the "functional status score" employed in Ayele [51], for example, measures activity, not quality—the two are not the same and should not be confused. Similarly, while Winch [49] and Omer [47] relied on patient report to assess the impact of CHWs, more rigorous, direct assessments of performance—like the service tests used by Carrón [50] and Gautham [42]—paired with biological measures would have been preferable. In future, trials

For peer review only

should provide information about implementation fidelity. The lack of consensus building efforts on metrics to CHW performance improvement is a hindrance to the accumulation of knowledge in this area.

Additionally, these interventions may have produced a Hawthorne effect. A number of studies documented that existing CHWs were under-supported in their work, and the intervention is likely to have been a motivational factor [42, 43ABC]. Moreover, the novelty—and in turn the effects—of such inputs may wear off over time [45ABC]. Future studies would do well to adjust for the attention effect and assess the long-term effects of such interventions.

Ultimately, for health tasks where CHW delivery demonstrate benefits, the focus of new research ought to shift from assessing efficacy to optimising public health impact [24]. This review found few trials on each of the hypothesised drivers of performance; component selection experiments on a greater range of interventions to improve performance ought to be conducted and used to refine existing theories of CHW performance [54, 55].

External validity and implications for practice

Several factors limit the applicability of these results [56]. Diverse outcomes across studies meant there were few data for each outcome of interest and multiple studies [42, 43] were underpowered to detect meaningful clinical difference. Because no studies reported adverse effects it is difficult to draw conclusions regarding the trade-off between benefits and harms [57].

Similarly, trials often failed to report relevant information regarding context. This is true both of the cultural context in which CHWs operate and CHW program components (e.g. recruitment, incentivisation etc.) that remained unmodified in the interventions.

For peer review only

The following factors about a potential implementation setting ought to be considered by decision-makers when assessing potential program-level performance improvement interventions [10, 57]:

1. Whether there are important differences in health system arrangements that may alter the feasibility and acceptability of the given intervention. For example, in Zambia [52] CHWs are part of the civil service and so eligible for government sponsored "in-service training." In countries where CHW programs are less well resourced it would not be possible to emphasise career advancement possibilities during recruitment.
2. Whether there are important differences in environmental conditions or other on-the-ground realities between where the studies were conducted and the implementation setting. Equipping CHWs with mobile phone-based guidance [42, 44] would not be possible in remote areas without mobile phone reception. Similarly, a supervision system premised on following-up with underperforming CHWs [45ABC], is impossible without a system to track CHW performance in the first place.
3. Whether there are important differences in the baseline health conditions between where the studies were done and the implementation setting. For example, tailoring of incentives [46] worked well for a program in which CHWs had a single categorical objective that only required one repetitive behaviour. Where CHWs have more complex objectives that require multiple behaviours performance-based incentives may distort behaviour in unintended and undesirable ways [58]—e.g. in Bossuroy [53] TB defaulting increased for patients of CHWs incentivised for case detection compared with patients of CHWs receiving a fixed salary.

For peer review only

4. Whether studies from which evidence was drawn were conducted in settings with similar social norms. Information about, for example, whether male CHWs in the Ayele [51] study faced cultural barriers working with women in the home etc. are examples of norms and attitudes that might have shaped intervention effects and so would be relevant to interpreting trial outcomes or exploring differential effects across settings.
5. Whether sufficient resources exist to implement the proposed clinical and/or managerial support for CHWs to a high standard and maintain that standard over time. The extent to which these proposed performance improvement interventions depend on the presence of other programmatic or contextual preconditions is not clear from the included studies—e.g. one study indicated that equipment may not be as valuable in the absence of adequate training [44].

Limitations

Firstly, RCTs on CHWs remain poorly indexed in electronic databases; it may therefore be possible that, despite the very careful and extensive searches and expert consultations, some relevant RCTs were not identified. Secondly, there is no widely accepted definition for this cadre of health worker. Though we have carefully specified the inclusion criteria in keeping with previous studies, there is a need for a clear taxonomy of CHW supports and activities. Thirdly, poor reporting and nonresponse from contacted authors meant the reviewers were unable to obtain some relevant missing data on methodological characteristics, clinical characteristics, and outcomes for some studies. Finally, though randomised trials have a unique value in intervention optimisation, exclusion of all non-randomised studies without further consideration places a zero weighting on such evidence. Updates of the systematic review should consider including rigorous non-experimental evidence.

For peer review only

Nevertheless, key data were available and rigorous methods have been applied throughout, including risk of bias assessments and evaluation of the quality of the evidence.

Previous Reviews

Between the original search and the update of this review, a systematic review on intervention design factors that influence CHW performance in LMICs was published [59]. Our more sensitive search strategy and focus on trials from which causal inferences can be made are important differences in approach.

CONCLUSION

Variations in recruitment, supervision, incentivisation, and equipment may improve CHW performance. Practitioners should, however, assess the relevance and feasibility of these strategies in their health setting prior to implementation and researchers should consider conducting component selection experiments on a greater range of interventions to improve performance. Nonetheless, mounting pressure to meet ambitious new international health goals and avoid repeating mistakes of the past underscore the timeliness and relevance of these findings.

Contributors: MB developed the study protocol with advice from PM. MB conducted the search. MB checked eligibility criteria, conducted data extraction and drafted the manuscript. All authors commented on the manuscript.

Acknowledgements: Sincere thanks to Calla Glavin for her help with study screening and data extraction.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests: None declared.

Data sharing statement: No additional data are available.

For peer review only

REFERENCES

1. World Health Organization, *Declaration of Alma-Ata*, in *International Conference on Primary Health Care*. 1978, World Health Organization: Alma-Ata, USSR.
2. Berman, P.A., D.R. Gwatkin, and S.E. Burger, *Community-Based Health-Workers - Head-Start or False Start Towards Health for All.* Social Science & Medicine, 1987. **25**(5): p. 443-459.
3. Frankel, S., *The community health worker: effective programmes for developing countries*. 1992: Oxford University Press.
4. Heggenhougen, K., et al., *Community health workers: the Tanzanian experience.* 1987: Oxford University Press.
5. Walt, G., *CHWs: are national programmes in crisis?* Health Policy and Planning, 1988. **3**(1): p. 1-21.
6. Gilson, L., et al., *National community health worker programs: how can they be strengthened?* Journal of public health policy, 1989: p. 518-532.
7. Walt, G., *Community health workers in national programmes: just another pair of hands?* 1990, Milton Keynes, UK: Open University Press.
8. Kredo, T., et al., *Task shifting from doctors to non-doctors for initiation and maintenance of antiretroviral therapy.* Cochrane Database of Systematic Reviews, 2014. **7**: p. CD007331.
9. Lassi, Z.S., B.A. Haider, and Z.A. Bhutta, *Community-based intervention packages for reducing maternal and neonatal morbidity and mortality and improving neonatal outcomes.* Cochrane Database of Systematic Reviews, 2010. **11**.
10. Lewin, S., et al., *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, 2010(3).
11. Okwundu, C.I., et al., *Home- or community-based programmes for treating malaria.* Cochrane Database of Systematic Reviews, 2013. **5**: p. CD009527.
12. United Nations. *The Millennium Development Goals Report 2015*. 2015 [cited 2016 12 June]; Available from: http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf.
13. Nullis-Kapp, C., *Health worker shortage could derail development goals.* Bulletin of the World Health Organization, 2005. **83**(1): p. 5-6.
14. Kruk, M.E., et al., *What is a resilient health system? Lessons from Ebola.* Lancet, 2015. **385**(9980): p. 1910-1912.
15. World Health Organization. *The Addis Ababa Declaration on task shifting.* in *The first international conference on task shifting*. 2008. Addis Abbaba: Health Systems and Services, WHO.
16. Campbell, J., et al., *A universal truth: No health without a workforce: Third global forum on human resources for health, Recife, Brazil.* Global Health Workforce Alliance and World Health Organization, Geneva, 2013.
17. Moszynski, P., *One billion people are affected by global shortage of healthcare workers.* British Medical Journal, 2011. **342**: p. d696.
18. Ghani, R., *40 million children have no access to the most basic healthcare services.* British Medical Journal, 2011. **343**: p. d4667.

For peer review only

19. Singh, P. and J.D. Sachs, *1 million community health workers in sub-Saharan Africa by 2015.* Lancet, 2013. **382**(9889): p. 363-365.
20. Munos, M., et al., *Independent Evaluation of the Rapid Scale-Up Program to Reduce Under-Five Mortality in Burkina Faso.* American Journal of Tropical Medicine & Hygiene, 2016. **94**(3): p. 584-95.
21. Amouzou, A., et al., *Effects of the integrated Community Case Management of Childhood Illness Strategy on Child Mortality in Ethiopia: A Cluster Randomized Trial.* American Journal of Tropical Medicine & Hygiene, 2016. **94**(3): p. 596-604.
22. Amouzou, A., et al., *Independent Evaluation of the integrated Community Case Management of Childhood Illness Strategy in Malawi Using a National Evaluation Platform Design.* American Journal of Tropical Medicine & Hygiene, 2016. **94**(3): p. 574-83.
23. Borrelli, B., *The Assessment, Monitoring, and Enhancement of Treatment Fidelity In Public Health Clinical Trials.* J Public Health Dent, 2011. **71**(s1): p. S52-S63.
24. Collins, L.M., et al., *The multiphase optimization strategy for engineering effective tobacco use interventions.* Ann Behav Med, 2011. **41**(2): p. 208-26.
25. Hasson, H., *Systematic evaluation of implementation fidelity of complex interventions in health and social care.* Implementation Science, 2010. **5**(1): p. 1.
26. Basch, C.E., et al., *Avoiding type III errors in health education program evaluations: a case study.* Health Education & Behavior, 1985. **12**(3): p. 315-331.
27. Alam, K., J.A.M. Khan, and D.G. Walker, *Impact of dropout of female volunteer community health workers: An exploration in Dhaka urban slums.* BMC Health Services Research, 2012. **12**.
28. Olang'o, C.O., I.K. Nyamongo, and J. Aagaard-Hansen, *Staff attrition among community health workers in home-based care programmes for people living with HIV and AIDS in western Kenya.* Health Policy, 2010. **97**(2-3): p. 232-237.
29. Medhanyie, A., et al., *Knowledge and performance of the Ethiopian health extension workers on antenatal and delivery care: a cross-sectional study.* Human Resources for Health, 2012. **10**.
30. Puchalski Ritchie, L.M., et al., *Evaluation of lay health workers' needs to effectively support anti-tuberculosis treatment adherence in Malawi.* Int J Tuberc Lung Dis, 2012. **16**(11): p. 1492-7.
31. Standing, H. and A.M.R. Chowdhury, *Producing effective knowledge agents in a pluralistic environment: What future for community health workers?* Social Science & Medicine, 2008. **66**(10): p. 2096-2107.
32. World Health Organization, *Community health workers: pillars for health for all. Report of the Interregional conference, Yaoundé, Cameroon, 1-5 December 1986*. 1987: Unpublished document SHS/CIH/87.2.
33. Khan, K.S., et al., *Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews*. 2001: NHS Centre for Reviews and Dissemination.
34. McAuley, L. and C. Ramsay. *Cochrane Effective Practice and Organisation of Care Review Group (EPOC) Data Collection Checklist*. 2002 February 4, 2014]; Available from:

For peer review only

https://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/datacollectionchecklist.pdf.

35. Hoffmann, T.C., et al., *Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide.* British Medical Journal, 2014. **348**.
36. Montgomery, P., et al., *The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses.* Journal of Clinical Epidemiology, 2013. **66**(8): p. 874-882.
37. Higgins, J.P.T. and S. Green, *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. 2011: The Cochrane Collaboration.
38. Guyatt, G.H., et al., *Rating quality of evidence and strength of recommendations: GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.* British Medical Journal, 2008. **336**(7650): p. 924.
39. Rehfuess, E.A. and E.A. Akl, *Current experience with applying the GRADE approach to public health interventions: an empirical study.* BMC Public Health, 2013. **13**(1): p. 9.
40. Waters, E., et al., *Essential components of public health evidence reviews: capturing intervention complexity, implementation, economics and equity.* Journal of Public Health, 2011. **33**(3): p. 462-465.
41. Movsisyan, A., G. Melendez-Torres, and P. Montgomery, *Users identified challenges in applying GRADE to complex interventions and suggested an extension to GRADE.* Journal of Clinical Epidemiology, 2015.
42. Gautham, M., M.S. Iyengar, and C.W. Johnson, *Mobile phone-based clinical guidance for rural health providers in India.* Health Informatics J, 2015. **21**(4): p. 253-66.
43. Derenzi, B., et al., *Improving Community Health Worker Performance Through Automated SMS.* ACM International Conference Proceeding Series, 2012: p. 25-34
44. Ginsburg, O.M., et al., *An mHealth model to increase clinic attendance for breast symptoms in rural Bangladesh: can bridging the digital divide help close the cancer divide?* Oncologist, 2014. **19**(2): p. 177-85.
45. Bailey, J.E. and D.W. Coombs *Effectiveness of an Indonesian model for rapid training of Guatemalan health workers in diarrhea case management*. Journal of Community Health, 1996. **21**, 269-76.
46. Andreoni, J., et al., *Using Preference Estimates To Customize Incentives: An Application To Polio Vaccination Drives In Pakistan*, in *NBER Working Paper Series*. 2016.
47. Omer, K., et al. *Evidence-based training of frontline health workers for door-to-door health promotion: a pilot randomized controlled cluster trial with Lady Health Workers in Sindh Province, Pakistan*. Patient Educ Couns, 2008. **72**, 178-85 DOI: 10.1016/j.pec.2008.02.018.
48. Chang, L.W., et al., *Impact of a mHealth Intervention for Peer Health Workers on AIDS Care in Rural Uganda: A Mixed Methods Evaluation of a Cluster-Randomized Trial.* AIDS and Behavior, 2011. **15**(8): p. 1776-1784.
49. Winch, P., et al., *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, 2003. **97**(5): p. 481-490.

For peer review only

50. Carrón, J., M. Melian, and F. Leon, *Final report of an operations research project:" Developing Tools of Low-Cost Use to Improve the Quality of Care of Rural CBD in Paraguay" Subcontract CI91. 96A January 1 1992-July 31 1994. Participating agencies: Centro Paraguayo de Estudios de Poblacion (CEPEP) the Population Council/INOPAL II*. 1994.
51. Ayele, F., A. Desta, and C. Larson, *The Functional Status Of Community-Health Agents - A Trial Of Refresher Courses And Regular Supervision.* Health Policy and Planning, 1993. **8**(4): p. 379-384.
52. Ashraf, N., O. Bandiera, and S.S. Lee, *Do-Gooders And Go-Getters: Career Incentives, Selection, And Performance In Public Service Delivery*. 2015.
53. Bossuroy, T., C. Delavallade, and V. Pons, *Fighting tuberculosis through community based counsellors: a randomized evaluation of performance based incentives in India*, in *3ie Grantee Final Report*. 2016, International Initiative for Impact Evaluation (3ie): New Delhi.
54. Kok, M.C., et al., *Which intervention design factors influence performance of community health workers in low- and middle-income countries? A systematic review.* Health Policy and Planning, 2014: p. [21] p.
55. Naimoli, J.F., et al., *A Community Health Worker "logic model": towards a theory of enhanced performance in low- and middle-income countries.* Human Resources for Health, 2014. **12**: p. 56.
56. Schünemann, H.J., A. Fretheim, and A.D. Oxman, *Improving the use of research evidence in guideline development: 13. Applicability, transferability and adaptation.* Health research policy and systems, 2006. **4**(1): p. 25.
57. Lavis, J.N., et al., *SUPPORT Tools for evidence-informed health Policymaking (STP) 16: Using research evidence in balancing the pros and cons of policies.* Health Research Policy and Systems, 2009. **7**(Suppl 1): p. S16.
58. Benabou, R. and J. Tirole, *Intrinsic and extrinsic motivation.* The Review of Economic Studies, 2003. **70**(3): p. 489-520.
59. Kok, M.C., et al., *Which intervention design factors influence performance of community health workers in low- and middle-income countries? A systematic review.* Health Policy and Planning, 2015. **30**(9): p. 1207-27.

For peer review only

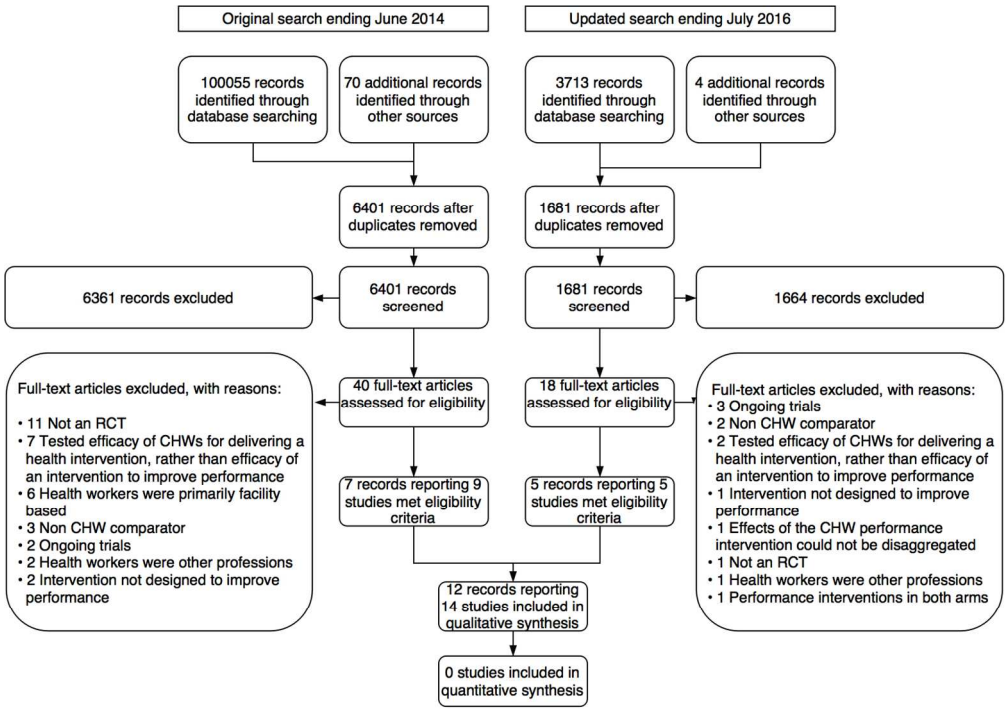


Figure 1 PRISMA study flow diagram

135x95mm (300 x 300 DPI)

For peer review only

Table of Contents

Appendix A: Protocol (Finalised April 2014) 1
Appendix B: Low and Middle-Income Countries 13
Appendix C: Details on the Search Strategy 13
Appendix D: Data Extraction Form 22
Appendix E: Characteristics of Included Studies 27
Appendix F: Characteristics of Excluded, Ongoing & Studies Awaiting Classification 49
Appendix G: Risk of bias of Included Studies 55
Appendix H: Cost Data 60
Appendix I: PRISMA Checklist 64

Appendix A: Protocol (Finalised April 2014)

The following protocol was reviewed by a panel of systematic review experts from the Centre for Evidence Based Intervention, University of Oxford in April 2014.

I. PICO QUESTION

Do interventions for improving the performance of community health workers, compared with treatment as usual, result in better primary healthcare outcomes in low- and middle-income countries?

II. BACKGROUND/STATE OF THE EVIDENCE

Description of the condition

Though widely promoted as a means to provide primary healthcare in resource poor settings as early as the 1978 Alma-Ata Declaration, community health workers (CHWs) have recently become the subject of renewed interest and debate in low and middle-income countries (LMICs) facing a growing human resource crisis (Gilmore &

For peer review only

McAulife, 2013; Lehmann & Sanders, 2007, WHO, 1978). The WHO estimates the world is currently short 7.2 million healthcare workers and that this deficiency is expected to grow to 12.9 million by the year 2035 (WHO/GHWA, 2013). This phenomenon is particularly severe in sub-Saharan Africa, which bears an estimated 24% of the world's burden of illness, yet has only 3% of the health care workers and 1% of the financial resources for health care worldwide (Anyangwe & Mtonga, 2007). Most acute in rural areas, personnel shortages mean that over one billion people around the world go their entire lives without seeing a health worker (Ghani, 2011; Moszynski, 2011). This lack of human resources for health (HRH) has significantly impeded progress toward the realisation of health-related Millennium Development Goals, namely to reduce child mortality (Goal 4), improve maternal health (Goal 5), and combat HIV and AIDS, malaria and other diseases (Goal 6) (Nullis-Kapp, 2005). CHWs have been proposed as a way to fill that gap by extending services to hard-to-reach populations and remote areas. Given mounting pressure to meet Millennium Development Goals, proposals to establish lay health workers—to whom simple medical procedures can be "task shifted" from doctors—have taken on revived urgency and prominence (Singh & Sachs, 2013; WHO/UNAIDS/PEPFAR, 2008).

Systematic reviews have concluded that CHWs are effective at delivering health services as diverse as immunization, case management of acute respiratory infections, malaria prophylaxis, and preventive interventions for maternal and child health (Christopher et al., 2011; Lehman & Sanders, 2007; Lewin et al. 2010). The reviews found that CHWs had a significant impact on reducing maternal, neonatal and child mortality in resource poor settings and increased coverage of health services when

For peer review only

compared to alternative modes of service delivery organisation (Gilmore & McAulife, 2013; Haines, 2007; Lassi et al., 2010). Although the reviews are of mixed quality—not all published protocols, assessed the quality of the sources compiled, or clearly delineated the data extraction process—the results from even the most exacting among them, a Cochrane review based solely on evidence from RCTs assessed for quality, indicated that CHWs "provide promising benefits in promoting immunisation uptake and breastfeeding, improving TB treatment outcomes, and reducing child morbidity and mortality when compared to usual care" (Lewin et al. 2010, pg. 2).

While these reviews examined the effects of CHWs compared to treatment as usual or other interventions, no systematic review has yet addressed the relative effectiveness of different types of CHW program designs. The question of how to optimise CHW performance—via different intervention components (e.g. approaches to recruitment, training, supervision, incentivization or other factors)—remains open (Lassi et al., 2010; Swider 2002).

Why it is important to do this review

While reviews have thus far isolated principles (e.g. "competent supervision") associated with CHW performance, none have identified actionable strategies to improve performance (e.g. "weekly clinical supervision by a nurse composed of supply audits, patient checks and on-the-job training"). This is the first systematic review to consider the effectiveness of interventions to improve the performance of community health workers.

III. OBJECTIVES

For peer review only

To assess the effects of interventions to improve the performance of community health workers for community-based primary health care outcomes in low income countries.

IV. METHODS

Criteria for considering studies for this review

Types of Studies

Because health workforce outcomes are empirically testable and workplace interventions possible to randomise, randomised controlled trials, cluster-randomised trials, and factorial trials will be eligible for inclusion.

Types of Participants

Community Health Workers:

Several definitions for and variants of the term community health worker have been employed in the literature; the roles and responsibilities of CHWs vary greatly depending on patients' access to facility-based care and the existence of other cadres of health workers. As Perry & Zulliger (2012) have noted, some CHWs have only a few days of training, while others have six months or more; some CHW cadres are paid salaries, others are volunteers; some are generalists, while others perform a narrowly defined set of interventions specific to one disease.

For the purpose of this review, definitions from Lewin et al. (2010) and the WHO (1987, 2007) were built on to define community health workers as any lay health workers who:

- Live and work in the community they serve
- Are part of the formal health system in a low- or middle-income country

For peer review only

- Perform interventions related to healthcare delivery
- Have received organised training in said interventions but have no formal or paraprofessional certification or tertiary education degree.

Although CHWs have been used in high-income countries, differences in healthcare budgets, education levels, access to care, and disease profile make it unproductive to include such actors in the comparison.

Recipients of care:

Aside from residence in a low- or middle-income country, there are no restrictions on the types of patients for whom data will be extracted.

Types of Interventions

Studies that evaluate any intervention designed to improve the performance of CHWs will be included. Studies will only be included where the primary objective, or one of the primary objectives, is to determine the efficacy or effectiveness of an intervention to improve the performance of CHWs. This includes studies like 'head-to-head' comparisons of different CHW interventions that have to date been excluded from reviews of CHW effectiveness (e.g. Lewin et al., 2010). Studies designed to examine the efficacy of a particular CHW activity (e.g. zinc tablets for diarrhoea) as opposed to the relative effectiveness of different ways of supporting CHWs to deliver said activity, will be excluded. Interventions of any duration or follow up will be included.

A study will be included if the intervention applied was adequately described (i.e. described in sufficient detail to understand the key steps undertaken) in the paper, in an appendix, or by the author in a personal communication.

Types of Outcome Measures

For peer review only

A number of processes and outcomes might be affected by interventions that aim to improve the performance of CHWs. Reported outcomes will be extracted and categorised as follows:

- Primary:
  - Objective physical outcomes for patients
  - Subjective health outcomes for patients
- Secondary
  - Utilisation indicators (e.g. mean number of consultations/month)
  - Quality of care outcomes (e.g. accuracy of diagnosis)
  - CHW retention

Search methods for identification of studies

Electronic Searches

The following strategy was used to search MEDLINE and was adapted for other databases:

1. ((community or village) adj2 (agent$ or aid$ or promot$ or OR mobili?er$ distribut?r$ or worker$)).ti,ot,ab,kf.
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliary$ or extension) adj2 (worker$ or volunteer$)).ti,ot,ab,kf.
3. (accompanier$ or accompagnateur$ or activista$ OR animatrice$ OR brigadista$ or kader$ or monitora$ or promotora$ or sevika$ or fhw$ or chw$ or lhw$ or vhw$ OR chv$ or "shastho shebika" or "shasto karmis" or anganwadi$ or

For peer review only

"barefoot doctor" OR "agente comunitario de salud" or "agente communitario de saude").ti,ot,ab,kf.

4. (performance or effective$ or skill$).ti,ot,ab,kf.
5. 1 or 2 or 3
6. 5 and 6
7. randomized controlled trial.pt.
8. controlled clinical trial.pt.
9. randomized.ab.
10. placebo.ab.
11. clinical trials as topic.sh.
12. randomly.ab.
13. trial.ti.
14. 7 or 8 or 9 or 10 or 11 or 12 or 13
15. 6 and 14

Databases searched:

- AMED, Ovid
- BiblioMap - EPPI-Centre database of health promotion research
- British Library for Development Studies at IDS
- CINAHL, Ebsco
- Cochrane Central Register of Controlled Trials (CENTRAL)
- Database of Promoting Health Effectiveness Reviews (DoPHER)
- EMBASE, Ovid

For peer review only

- Global Health
- IDEAS database of unpublished working papers
- MEDLINE In-Process & Other Non-Indexed Citations
- MEDLINE, Ovid
- POPLINE
- Pubmed
- Social Science Citation Index/Web of Science
- WHOLIS
- World Bank's JOLIS search engine

Searching other resources

With the use of a pre-determined search strategy, studies will be identified through:

1. Systematic search of electronic databases using a pre-defined search strategy, described above.
2. Handsearch of relevant journals, search of trial registries for on-going research (e.g. ClinicalTrials.gov), and search of grey literature databases (e.g. Open GREY)
3. Review of websites and via direct contact with local agencies, research institutions, relevant government departments, and international donors and multilateral agencies involved in HIV care, treatment, and prevention. These organisations will include GHWA, UNICEF, WHO, & USAID.
4. Direct contact with experts in the field for unpublished studies.

For peer review only

5. Review of references from included studies and relevant articles to identify other studies that met the inclusion criteria.

Data Collection and analysis

Selection of studies

Two authors will independently screen titles, abstracts, and full text articles. Included articles will be determined by consensus.

Data extraction and management

Reviewers will extract data on study design, participants, interventions, outcomes and methodological quality, using a modified version of the Cochrane EPOC Group data collection checklist. Data will be managed using Excel.

Assessment of risk of bias in included studies

The Cochrane Collaboration's tool for evaluating the risk of bias will be used to assess individual studies. The quality of evidence across a body of evidence will be assessed using the GRADE approach.

Measures of treatment effect

Statistical analysis will be carried out using the Review Manager software (RevMan 2012). Measures of treatment effect will be standardised (dichotomous: risk ratios, continuous: standardised mean difference) where possible or, if not, presented as reported by the trial authors with 95% CIs and tests of statistical significance where available.

Unit of analysis issues

Cluster-randomised trials will be verified to ensure appropriate analysis has been done (i.e. one that adjusts for clustering in calculating confidence intervals or P-values).

For peer review only

If such an analysis has not been done, the necessary data (intracluster correlation coefficients - ICCs) will be extracted if possible or obtained from the investigators and results will be re-analysed. If this is not possible, point estimates, but not the reported confidence intervals or P-values, will be reported. If there are sufficiently similar studies to conduct meta-analyses, a sensitivity analysis will be undertaken using imputed ICCs based on data from the other included studies and other studies with comparable outcome measures (Campbell, 2000).

For cross-over trials, we will extract and analyse data only from the period before the crossover. For factorial studies, we will include all comparisons.

Dealing with missing data

Authors of each of the included studies will be contacted to obtain missing statistics (i.e. unreported data). Missing data due to participant attrition will be dealt with by conducting sensitivity analyses to investigate attrition as a source of heterogeneity and possible bias.

Assessment of heterogeneity

If meta-analyses is conducted, the extent of heterogeneity in results across comparable studies will be assessed using forest plots, the I^2 statistic and the Chi^2 test.

Assessment of reporting biases

Selective outcome reporting will be assessed using the approach described in Chapter 8 of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2008). Publication bias will be assessed using funnel plots (plotting effect size against standard error) to detect potential bias. Because asymmetry can be due to clinical and

For peer review only

methodological heterogeneity in addition to publication bias, these sources of heterogeneity will also be examined as possible explanations.

Data synthesis

Given the question asked in the review is broad with respect to the type of intervention and outcome considered, a random effects analysis will be conducted if possible. If no metaanalysis is conducted, a narrative synthesis will be conducted.

Subgroup analysis and investigation of heterogeneity

No subgroup analysis will be performed.

Sensitivity analysis

A sensitivity analyses will be performed by excluding studies with a high risk of bias for any outcome for which more than one comparable study with studies with a low or moderate risk of bias.

Differences Between Protocol and Review (September 2016)

- Four additional databases were added in July of 2016 to better capture the international development literature (3ie Impact Evaluation Database, IPA Database, J-PAL Evaluations Database, World Bank Independent Evaluation Group)

References for Appendix A

Anyangwe S, Mtonga C. (2007) Inequities in the global health workforce: The greatest impediment to health in sub-Saharan Africa. *International Journal of Environmental Research and Public Health*, 4:93-100

Christopher, J. B., Le May, A., Lewin, S., & Ross, D. a. (2011). Thirty years after Alma-Ata: a systematic review of the impact of community health workers delivering curative interventions against malaria, pneumonia and diarrhoea on child mortality and morbidity in sub-Saharan Africa. *Human resources for health*, *9*(1), 27. doi:10.1186/1478-4491-9-27

For peer review only

Ghani, R. (2011). 40 Million children have no access to the most basic healthcare services. *BMJ (Clinical research ed.)*, *343*(jul22_1), d4667. doi:10.1136/bmj.d4667

Gilmore, B., & McAuliffe, E. (2013). Effectiveness of community health workers delivering preventive interventions for maternal and child health in low- and middle-income countries: a systematic review. *BMC public health*, *13*(1), 847. doi:10.1186/1471-2458-13-847

Lassi, Z., Haider, B., & Bhutta, Z. (2010). Community-based intervention packages for reducing maternal and neonatal morbidity and mortality and improving neonatal outcomes (Review). *Public health nursing (Boston, Mass.)*, *28*(3), 246–8. doi:10.1111/j.1525-1446.2011.00953.x

Lehmann, U., & Sanders, D. (2007). *Community health workers: What do we know about them?* (pp. i–34). Geneva.

Lewin, S., Munabi-Babigumira, S., Glenton, C., Daniels, K., Bosch-Capblanch, X., VanWyk, B. (2010). Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases (Review). *Public health nursing (Boston, Mass.)*, *28*(3), 243–5. doi:10.1111/j.1525-1446.2011.00950.x

Moszynski, P. (2011) One billion people are affected by global shortage of healthcare workers. *BMJ* 2011;342:d696.

Nullis-Kapp, C. (2005). Health worker shortage could derail development goals. *Bulletin of the World Health Organization, 83*(1), 5-6.

Perry, H., & Zulliger, R. (2012). *How Effective Are Community Health Workers?* MDG Health Alliance.

Review Manager (RevMan). 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012.

Singh, P., & Sachs, J. D. (2013). 1 million community health workers in sub-Saharan Africa by 2015. *Lancet*, *382*(9889), 363–5. doi:10.1016/S0140-6736(12)62002-9

Swider, S. M. (2002). Outcome effectiveness of community health workers: an integrative literature review. *Public Health Nursing (Boston, Mass.)*, *19*(1), 11–20. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11841678

WHO (2007). Task shifting: rational redistribution of tasks among health workforce teams. Geneva: World Health Organization.

WHO/Global Health Workforce Alliance (2013) A Universal Truth: No Health Without a Workforce. Geneva [http://www.who.int/workforcealliance/knowledge/resources/hrhreport2013/en/].

WHO/UNAIDS/PEPFAR (2008) Task Shifting: Global Recommendations and Guidelines. Geneva [http://www.who.int/healthsystems/task_shifting/en/index.html].

WHO (1978). Declaration of Alma-Ata. Paper presented at the International Conference on Primary Health Care, Alma-Ata, USSR.

WHO (1987). Community Health Workers: Working Document for the WHO Study Group. Geneva, Switzerland: WHO.

For peer review only

Appendix B: Low and Middle-Income Countries

Based on World Bank Country and Lending Groups:

Upper-middle-income economies
Albania, Algeria, Algeria, American Samoa, Angola, Azerbaijan, Belarus, Belize, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, China, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, Fiji, Gabon, Grenada, Iran, Islamic Rep., Iraq, Jamaica, Jordan, Kazakhstan, Lebanon, Libya, Macedonia, Malaysia, Maldives, Marshall Islands, Mauritius, Mexico, Mongolia, Montenegro, Namibia, Palau, Panama, Paraguay, Peru, Romania, Serbia, South Africa, St. Lucia, St. Vincent and the Grenadines, Suriname, Thailand, Tonga, Tunisia, Turkey, Turkmenistan, Tuvalu.

Lower-middle-income economies
Armenia, Bangladesh, Bhutan, Bolivia, Cabo Verde, Cameroon, Congo, Rep., Côte d'Ivoire, Djibouti, Egypt, El Salvador, Georgia, Ghana, Guatemala, Guyana, Honduras, India, Indonesia, Kenya, Kiribati, Kosovo, Kyrgyz Republic, Lao PDR, Lesotho, Mauritania, Micronesia, Moldova, Morocco, Myanmar, Nicaragua, Nigeria, Pakistan, Papua New Guinea, Philippines, Samoa, São Tomé and Principe, Senegal, Solomon Islands, Sri Lanka, Sudan, Swaziland, Syrian Arab Republic, Tajikistan, Timor-Leste, Ukraine, Uzbekistan, Vanuatu, Vietnam, West Bank and Gaza, Yemen, Zambia

Low-income economies
Afghanistan, Benin, Burkina Faso, Burundi, Cambodia, Central African Republic, Chad, Comoros, Dem. Rep Congo, Eritrea, Ethiopia, The Gambia, Guinea, Guinea-Bisau, Haiti, Dem Rep. Korea, Liberia, Madagascar, Malawi, Mali, Mozambique, Nepal, Niger, Rwanda, Sierra Leone, Somalia, South Sudan, Tanzania, Togo, Uganda, Zimbabwe

Appendix C: Details on the Search Strategy

In addition to the electronic sources listed in the main text, studies were identified through:

1. Handsearch of (a) seven journals from 1960 (or first online issue) to May 2014 (*Health Policy and Planning*, *Journal of Community Health*, *Social Science and Medicine*, *Human Resources for Health*, *BMC Public Health, Global Health Action*, and *Global Public Health*) and (b) conference proceedings from the American Society of Tropical Medicine and Hygiene.

For peer review only

2. Review of websites of nineteen leading global public health research institutions, donors, implementers, and technical agencies: International Child Health Review Collaboration; World Health Organization; United Nations Children's Fund; United Nations Development Programme; United Nations Population Fund; Joint United Nations Programme on HIV/AIDS; United States Agency for International Development; UK Department for International Human Development Resource Centre; Partners in Health; the Global Fund to Fight AIDS, Tuberculosis and Malaria; Bill & Melinda Gates Foundation; Clinton Foundation; Centre for Global Development; University Research Co-Center for Human Services (URC-CHS); CARE, Frontline Health Workers Coalition; One Million Community Health Workers Campaign; Global Health Workforce Alliance; and CHW Central.
3. Search of grey literature databases relevant to public health: Open GREY, New York Academy of Medicine Grey Literature Report, BIOSIS, and Proquest Dissertations & Theses.
4. Search of trial registries for on-going research: WHO International Clinical Trials Registry Platform, the ISRCTN Register, and ClinicalTrials.gov.
5. Direct contact with study authors and leading experts in the field of community health to solicit potentially relevant unpublished papers, on-going research, and suggestions for other contacts (see Appendix 6a for full list).
6. Examination of the reference lists of related reviews and primary studies to identify other studies that meet the inclusion criteria.

Electronic search strategies are as follows:

AMED, Ovid

For peer review only

AMED (1985 to April May 2014) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 53 hits. The search was updated May 11, 2016 and produced 4 hits.

BiblioMap: EPPI-Centre database of health promotion research

The search strategy used for BiblioMap was as follows:

1. Community Health Agent
2. Community Health Aides
3. Community health promoter
4. Community mobilizer
5. Community drug distributor
6. Community health worker
7. Village health worker
8. Rural Health Worker
9. Lay Health Worker
10. Lady health worker
11. Nutrition worker
12. Frontline health worker
13. Barangay health worker
14. Basic health worker
15. Auxiliary health worker
16. health extension worker
17. community health volunteer
18. village health volunteer
19. accompanier
20. accompagnateur
21. activista
22. animatrice
23. brigadista
24. kader
25. promotora
26. monitora
27. sevika
28. fhw
29. lhw
30. chw
31. vhw
32. chv
33. shastho shebika
34. shasto karmis
35. anganwadi
36. barefoot doctor

For peer review only

37. agente comunitario de salud
38. agente communitario de saude
39. 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
40. performance
41. effectiveness
42. skill
43. 40 OR 41 OR 42
44. 39 AND 43

The search was conducted on June 30, 2014 and produced 8 hits. The search was updated May 11, 2016 and produced 0 hits.

British Library for Development Studies at IDS

The search strategy used for IDS was as follows: "community health workers"

The search was conducted on June 30, 2014 and produced 5 hits. The search was updated May 11, 2016 and produced 4 hits.

CINAHL, EBSCO

CINAHL, Cumulative Index to Nursing and Allied Health (1960 to May 2014) was searched via Ebsco using the following terms:

> (TI ((community OR village) N2 (agent* OR aid* OR promot* OR mobili?er* OR distribut?r* OR worker)) OR TI ((village OR rural OR lay OR lady OR nutrition OR frontline OR barangay OR basic OR auxiliar* OR extension) N2 (worker* OR volunteer*)) OR TI ((accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR monitora* OR promotora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente communitario de saude"))) AND AB ((performance or effective* or skill*))

The search was conducted on May 4, 2014 and produced 189 hits. The search was updated May 11, 2016 and produced 31 hits.

For peer review only

CENTRAL

The search strategy used for the Cochrane Central Register of Controlled Trials (Issue 5 of 12, May 2014) was as follows:

1. ((community or village) near/2 (agent* or aid* or promot* or mobili*er* or distribut*r* or worker*)):ti,ab
2. ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliar* or extension) near/2 (worker* or volunteer*)):ti,ab
3. (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or promotora* or monitora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "shastho shebika" or "shasto karmis" or anganwadi* or "barefoot doctor" or "agente comunitario de salud" or "agente communitario de saude"):ti,ab
4. (performance or effective* or skill*):ti,ab
5. #1 or #2 or #3
6. #4 and #5

(In: Trials)

The search was conducted on May 4, 2014 and produced 256 hits. The search was updated May 11, 2016 and produced 129 hits.

Database of Promoting Health Effectiveness Reviews (DoPHER) search strategy
The search strategy used for DoPHER was as follows:

1. Community Health Agent
2. Community Health Aides
3. Community health promoter
4. Community mobilizer
5. Community drug distributor
6. Community health worker
7. Village health worker
8. Rural Health Worker
9. Lay Health Worker
10. Lady health worker
11. nutrition worker
12. frontline health worker
13. Barangay health worker
14. basic health worker
15. Auxiliary health worker

For peer review only

16. health extension worker
17. community health volunteer
18. village health volunteer
19. accompanier
20. accompagnateur
21. activista
22. animatrice
23. brigadista
24. kader
25. promotora
26. monitora
27. sevika
28. fhw
29. lhw
30. chw
31. vhw
32. chv
33. shastho shebika
34. shasto karmis
35. anganwadi
36. barefoot doctor
37. agente comunitario de salud
38. agente communitario de saude
39. 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
40. performance
41. effectiveness
42. skill
43. 40 OR 41 OR 42
44. 39 AND 43

The search was conducted on June 30, 2014 and produced 8 hits. The search was updated May 11, 2016 and produced 0 hits.

EMBASE search strategy

EMBASE (1974 to 2014 Week 18) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 2526 hits. The search was updated May 11, 2016 and produced 796 hits.

Global Health search strategy

For peer review only

Global Health (1973 to 2014 Week 18) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 1341 hits. The search was updated May 11, 2016 and produced 450 hits.

IDEAS database of unpublished working papers

The search strategy used for the IDEAS database was as follows:

> "community health workers" performance
> Match: all; Words forms: exact; In: abstract

The search was conducted on June 30, 2014 and produced 4 hits. The search was updated May 11, 2016 and produced 4 hits.

Impact Evaluation Database (3ie)
Keyword: community health worker
Evaluation method: Randomised Control Trials (RCT)

The search was conducted on July 12, 2016 and produced 62 hits.

IPA (Innovations for Poverty Action) Database
Keyword: community
Program area: Health

The search was conducted on July 12, 2016 and produced 6 hits.

Databases within ISI Web of Science

Databases within ISI Web of Science (Science Citation Index Expanded, Social Sciences Citation Index, Conference Proceedings Citation Index – Science, Conference Proceedings Citation Index - Social Sciences & Humanities, Book Citation Index– Science, Book Citation Index– Social Sciences & Humanities) (1945-May 2014) were searched using the following strategy:

> TITLE: (((community OR village) NEAR/2 (agent* OR aid* OR promot* OR mobili*er* OR distribut*r* OR worker*))) OR TITLE: (((village OR rural OR lay

For peer review only

OR lady OR nutrition OR frontline OR barangay OR basic OR auxiliar* OR extension) NEAR/2 (worker* OR volunteer*))) OR TITLE: ((accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR promotora* OR monitora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente communitario de saude")) AND TITLE: ((performance or effective* or skill*))
Timespan: All years. Indexes**:** SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH,

The search was conducted on May 4, 2014 and produced 3047 hits. The search was updated May 11, 2016 and produced 657 hits.

J-PAL (Abdul Latif Jameel Poverty Action Lab) Evaluations Database
Keyword: Community worker
Sector: Health

The search was conducted on July 12, 2016 and produced 23 hits.

MEDLINE search strategy

MEDLINE (1946 to May 2014 Week 19) was searched via Ovid using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 2143 hits. The search was updated May 11, 2016 and produced 405 hits.

MEDLINE In-Process In-Process & Other Non-Indexed Citations search strategy
MEDLINE In-Process (1946 to May 2014 Week 19) was searched via OVID using the terms in the main body of the text. The search was conducted on May 4, 2014 and produced 251 hits. The search was updated May 11, 2016 and produced 734 hits.

POPLINE, K4Health

The search strategy used for POPLINE (1965-May 2014) was as follows:

All fields: effective* OR performance OR skill AND

Title: ((community or village) near/2 (agent* or aid* or promot* or mobili*er* or

For peer review only

distribut*r* or worker*)) OR ((village or rural or lay or lady or nutrition or frontline or barangay or basic or auxiliar* or extension) near/2 (worker* or volunteer*)) OR (accompanier* or accompagnateur* or activista* or animatrice* or brigadista* or kader* or promotora* or monitora* or sevika* or fhw* or chw* or lhw* or vhw* or chv* or "shastho shebika" or "shasto karmis" or anganwadi* or "barefoot doctor" or "agente comunitario de salud" or "agente communitario de saude")

The search was conducted on May 4, 2014 and produced 20 hits. The search was updated May 11, 2016 and produced 168 hits.

Pubmed

The search strategy used for Pubmed (1960-May 2014) was as follows:

((((("Community health agent" or "Community Health Aides" or "Community health promoter" or "Community mobilizer" or "Community drug distributor" or "community health worker" or "Village health worker"[Title/Abstract])) OR ("Rural Health Worker" or "Lay Health Worker" or "Lady health worker" or "nutrition worker" or "frontline health worker" or "Barangay health worker" or "basic health worker" or "Auxiliary health worker" or "health extension worker" or "community health volunteer" or "village health volunteer"[Title/Abstract])) OR (accompanier* OR accompagnateur* OR activista* OR animatrice* OR brigadista* OR kader* OR promotora* OR monitora* OR sevika* OR fhw* OR chw* OR lhw* OR vhw* OR chv* OR "shastho shebika" OR "shasto karmis" OR anganwadi* OR "barefoot doctor" OR "agente comunitario de salud" OR "agente communitario de saude"[Title/Abstract])) AND (performance or effective* or skill*[Title/Abstract]))NOT MEDLINE[sb])

The search was conducted on May 4, 2014 and produced 195 hits. The search was updated May 11, 2016 and produced 233 hits.

WHOLIS

WHOLIS (1948-May 2014) was searched via the Global Health Library the following terms: "community health workers" and performance.

For peer review only

The search was conducted on June 30, 2014 and produced 3 hits. The search was updated May 11, 2016 and produced 0 hits.

World Bank's JOLIS

JOLIS was searched with the following terms:

"community health workers" and performance

The search was conducted on June 30, 2014 and produced 6 hits. The search was updated May 11, 2016 and produced 0 hits.

World Bank Independent Evaluation Group

Advanced Search>>[all items]>>Systematic Reviews and Impact Evaluations>>Health, Nutrition and Population

The search was conducted on July 12, 2016 and produced 7 hits.

Appendix D: Data Extraction Form

STUDY ID (Author name, Year):	

VARIABLE	VALUE/ JUDGE-MENT	SOURCE	OPTIONS	VARIABLE TYPE
• INCLUSION CRITERIA				
Types of Studies				
Study design			• RCT • C-RCT • Factorial • ITS • Cohort • Case Control • CBA • Other (specify)	Nominal
Types of Participants				
Live and work in the catchment they serve.			• Yes • No • Unclear	Nominal

For peer review only

Are primarily based in said community			• Yes • No • Unclear	Nominal
Belong to of the formal health system in a low-income country			• Yes • No • Unclear	Nominal
Perform interventions related to healthcare delivery.			• Yes • No • Unclear	Nominal
Have received organised training in said interventions but have no formal or paraprofessional certification or tertiary education degree			• Yes • No • Unclear	Nominal
Types of Interventions				
Primary objective to determine the efficacy or effectiveness of an intervention to improve the performance of CHWs			• Yes • No • Unclear	Nominal
Primary objective to examine the efficacy of a particular CHW activity as opposed to the relative effectiveness of different ways of supporting CHWs to deliver said activity			• Yes • No • Unclear	Nominal
Type of intervention			• Recruitment • Training • Supervision • Incentives • Other (describe)	Nominal
Types of Outcomes				

For peer review only

List				Narrative
Decision				
Include or exclude			• Include • Exclude	Dichotomous
Reason for exclusion:				Narrative
•METHODS				
Study duration			• # months	Continuous
Date of study			• MM/YYYY-MM/YYY	
Unit of allocation			• Patient • Provider • Community • Other (specify) • Unclear	Nominal
•PARTICIPANTS				
Total number			• # intervention, # control	Continuous
Setting			• Rural • Urban • Mixed (specify)	Nominal
Country				Nominal
Local name for CHWs				Nominal
Definition of CHW				Narrative
Age			• Mean (years)	Continuous
Sex			• % Male and % Female	Continuous
Education level			• Some primary • Primary • Some secondary • Secondary • Tertiary	Ordinal
Services provided				Narrative
Patients served/CHW			• Mean	Continuous
Reimbursement			• Patient pays fee for service • Patient receives care for free	Nominal

For peer review only

•INTERVENTION DESCRIPTION				
Number of intervention groups			•#	Continuous
Description of Intervention (who delivered, format and timing of delivery)				Narrative
Description of control				Narrative
Integrity of intervention				
•OUTCOMES				
List outcomes (definition, unit of measurement)				Narrative
For each outcome, time points (i) collected (ii) reported				
Unit of analysis			• Patient • Provider • Community • Other (specify) • Not clear	Nominal
Adverse or unintended effects?				Narrative
•RESULTS				
For each intervention group:				
Number of participants allocated			• #	Continuous
Sample size			• #	Continuous
Missing participants			• #	Continuous
Summary data			• 2x2 OR Mean, SD	
•PROGRAM CHARACTERISTICS				
Recruitment				

For peer review only

Selected by			• Community • Health system • Both • Other (specify)	Nominal
Process			• Describe	Narrative
Training				
Amount			• <1 day • 1 day-1 week • 1 week - 3 months • > 6 months	Ordinal
Frequency			• Once • Weekly • Monthly • Quarterly • Yearly	Ordinal
Assessment			• Required (describe) • Not required	Dichotomous
Supervision				
Type			• Clinical • Peer • None	Nominal
Frequency			• None • Daily • Weekly • Monthly • Quarterly	Ordinal
Incentivisation				
Type			• Monetary (amount) • Non-monetary • None	Nominal
•MISCELLANEOUS				
Funding source				
Power calculation			• Done • Not clear • Not done	Nominal
Conclusions of study authors				Narrative

For peer review only

Equity considerations: evidence of differential impact on different parts of the population				Narrative
•Notes				

Appendix E: Characteristics of Included Studies

Andreoni 2016

Trial reg.	AEARCTR-0000417
Country	Pakistan
Methods	STUDY DESIGN: RCT (2x2 Factorial) UNIT OF RANDOMISATION: LHW
Population	LOCAL TERM FOR CHW: Lady Health Workers (LHWs) CHWs: 100% female, mean 10 years of education, age not specified PATIENTS: Children under five HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: (A): 85 CHWs, (B): 84 CHWs CONTROL SAMPLE SIZE: (A): 88 CHWs, (B): 80 CHWs
Interventions	OBJECTIVE: Use structural estimates of time preferences to customise incentives for a sample of polio vaccinators during a series of door-to-door immunization drives in Pakistan and experimentally evaluate the effect of matching contract terms to individual discounting patterns. The study investigates the effect of tailoring, but also whether tailoring is more or less effective depending on when allocations are made. INTERVENTION DESCRIPTION: The function of LHWs is to provide oral polio vaccine to children during government organized vaccination drives; LHWs are given a supply of oral vaccine and a neighbourhood map and are asked to travel door-to-door vaccinating children with a suggested target for vaccination. LHWs are offered a fixed bonus of 1000 rupees (around $10) for completing a total of V = 300 vaccinations over a two- day drive. Vaccinators set daily targets v1 and v2 corresponding to vaccinations on day 1 and day 2 of the drive, respectively. If either of the vaccination targets, v1 or v2, are not met, the bonus is not received. Vaccinators are randomly assigned an interest rate, R, such that a single vaccination that is allocated to day 2 reduces by R the number of vaccinations required on day 1. That is, v1 and v2 satisfy the intertemporal budget constraint. There were two intervention groups: (A) Contracts tailored to individual discounting patterns with advance allocation choice and (B) Contracts tailored to individual discounting patterns with immediate allocation choice. TRAINING DURATION (DAYS): 1 TRAINING DESCRIPTION: No description of baseline training. All CHWs received a two hour training on the smartphone vaccination application and its correct use. Independent of treatment status, LHWs were also trained on the intertemporal bonus contracts and the process by which allocations were made and submitted. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Daily

For peer review only

	SUPERVISION DESCRIPTION: At the end of each day of the drive, LHWs in each neighbourhood convened at a local clinic with their supervisor, and self-reported their vaccinations for the day. Though, in principle, a monitor could investigate the neighbourhood chalk markings, in practice, this system provided for virtually no monitoring. INCENTIVE: Monetary INCENTIVE DESCRIPTION: No explicit incentives for completing vaccinations were provided and LHWs received a fixed daily wage of 100 rupees (around $1). As part of the trial, LHWs were offered fixed bonus of 1000 rupees (around $10) for completing a total of V = 300 vaccinations over a two- day drive. RATIO CHW TO POPULATION: N.S. START DATE: September 26, 2014 END DATE: December 9, 2014 CONTROL DESCRIPTION: Prior to our project, LHWs self-reported achievement and no technology existed for monitoring vaccinations. Each vaccinator in the sample was provided a smart-phone, equipped with a reporting application that permits precise observation of when and where vaccinations are conducted. There were two control groups: (A) Untailored contracts and advance allocation choice and (B) Untailored contracts and immediate allocation choice.
Outcomes	PATIENTS: None CHWs: Closeness to the policy objective DURATION OF FOLLOW UP: Immediate
Funding source	International Growth Center (IGC), National Science Foundation, grant SES-1427355, and from the Department for International Development (DFID) Building Capacity for Research Evidence (BCURE) pilot program initiative
Notes	BASELINE INPUT: Smartphones, as described above

Risk of bias

Item	**Reviewers' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "We randomly provide half of LHWs with a tailored contract and half with a random contract" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear risk	Comment: Not specified.
Blinding of participants and personnel (performance bias)	Unclear risk	Comment: Not specified.
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Not specified.
Incomplete outcome data (attrition bias)	Low risk	Comment: None reported
Selective reporting (reporting bias)	Unclear risk	Comment: Outcomes listed in the protocol were embargoed so it is not possible to make a judgement

Ashraf 2015

Trial reg.	None
Country	Zambia
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: District CLUSTERS: 24 Intervention, 24 Control

For peer review only

Population	LOCAL TERM FOR CHW: Community Health Agent (CHA) CHWs: 30% female, 18-45 years (mean 27 years), minimum 12 years of education PATIENTS: Women and children under 5 years, mean 4.2 years education HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 149 CHWs CONTROL SAMPLE SIZE: 149 CHW
Interventions	OBJECTIVE: Identify the effect of advertising career incentives at the recruitment stage. INTERVENTION DESCRIPTION: Recruitment materials make career possibilities salient by highlighting that CHAs are part of the Ministry of Health's hierarchy and that this gives them access to a career path leading to higher-ranked positions such as nurse, clinical officer, and doctor. (Being part of the civil service, CHAs are eligible for "in- service training," meaning that they attend school as a serving officer and the government pays their tuition for all of their training.) Example line from the poster: "Become a community health worker to gain skills and boost your career!" TRAINING DURATION (DAYS): 365 TRAINING DESCRIPTION: Learn basic medical practices at a residential training. REFRESHER OR ONGOING TRAINING: Refresher training referenced but not described. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: 5% of home visits are audited; CHWs meet with their supervisors at an unspecified frequency. INCENTIVE: Monetary INCENTIVE DESCRIPTION: USD 290/month RATIO CHW TO POPULATION: N.S. START DATE: August, 2010 END DATE: January, 2014 CONTROL DESCRIPTION: Recruitment materials make salient benefits to the community, thus making the CHA position look similar to existing informal positions (e.g., village health workers, traditional birth attendants, barefoot doctors) that are common in these areas. Example line from the poster: "Want to serve your community? Become a community health worker!"
Outcomes	PATIENTS: Use of health services CHWs: Number of household visits completed over the study period, number of patients seen at health post, number of community meetings organised DURATION OF FOLLOW UP: 18 months
Funding source	IGC, JPAL Governance Initiative and HBS DFRD
Notes	BASELINE INPUT: CHAs, after a year of training, are required to devote 80% of their time (4 out of 5 working days per week) to household visits. The visits' main goals are to provide advice on women's health—including family planning, pregnancy, and postpartum care—and child health—including nutrition and immunizations. In addition, CHAs are expected to inspect the household and provide advice on health-related practices such as safe water practices, household waste management, sanitation, hygiene and ventilation. During visits, CHAs are also tasked with providing basic care to any sick persons and referring them to the health post as needed. In the remaining time, CHAs are expected to assist staff at the health post (the first-level health facility in rural Zambia) by seeing patients, assisting with antenatal care, and maintaining the facility. They are also supposed to organize community meetings such as health education talks at the health post and in schools.

Risk of bias

Item	Reviewers' judgement	Support for judgement

For peer review only

Random sequence generation (selection bias)	Unclear risk	Quote: "Random assignment of the 48 districts is stratified by province and average district-level educational attainment" Comment: Unclear how this was done.
Allocation concealment (selection bias)	Unclear risk	Comment: Not specified.
Blinding of participants and personnel (performance bias)	Unclear risk	Comment: Not specified.
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Not specified.
Incomplete outcome data (attrition bias)	High risk	Quote: "18% of CHAs drop out" after one year. "This attrition rate is balanced across treatments"
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol was identified
Recruitment bias	Low risk	Comment: CHWs were recruited after the study, but by district officials exposed to one treatment and unaware of the other
Baseline imbalance	Unclear risk	Comment: Not specified
Loss of clusters	Unclear risk	Comment: Not specified
Unit of analysis issues	Low risk	Quote: "standard errors clustered at the district level"

Ayele 1993

Trial reg.	None.
Country	Ethiopia
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: District CLUSTERS: Intervention: 1, Control: 1
Population	LOCAL TERM FOR CHW: Community health agents (CHAs) CHWs: 100% Male. Mean age 31.0 years. Mean schooling 5.0 years. PATIENTS: Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 50 CHWs, number of patients unspecified CONTROL SAMPLE SIZE: 50 CHWs, number of patients unspecified
Interventions	OBJECTIVE: Determine the effect of a refresher course and monthly supervision on the health service activities of CHWs INTERVENTION DESCRIPTION: Intervention group (N= 50 CHWs) received (1) Training: refresher course (2) Supervision: monthly INTERVENTION TYPE: Training-Supervision TRAINING DURATION (DAYS): 5 TRAINING DESCRIPTION: CHWs in the intervention district received (1) a five-day refresher course based on the "expected functions of the CHWs and other relevant responsibilities which they are asked to carry out in the provision of the health care to their communities" REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Monthly SUPERVISION DESCRIPTION: Educative supervision by health unit staff. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S.

For peer review only

	RATIO CHW TO POPULATION: N.S. START DATE: 06/1990 END DATE: 12/1990 CONTROL DESCRIPTION: CHWs (N = 52 CHWs) received nothing after baseline input (see notes). CHWs in control group had gone an average of 1.9 years since last refresher training and 30.8% had never received supervision.
Outcomes	PATIENTS: None. CHWs: (1) Composite functional status score (FSS) computed by summing the scores for the 13 CHW activities (2) 13 activity scores: (i) Outreach (days/mo.), (ii) Health education (sessions/mo.), (iii) Environmental health (activities/mo.), (iv) Maternal and child health (activities/mo.), (v) Expanded programme on immunization (activities/mo.), (vi) Births (# registered/mo.), (vii) Deaths (# registered/mo.), (viii) Curative (pts./mo.), (ix) Reports (#/mo.), (x) School health (activities/mo.), (xi) Home visits (#/mo.), (xii) Referrals (#/3mo), (xiii) Epidemic control measures (#/3mo). DURATION OF FOLLOW UP: Collected monthly for 6 months and reported at 3 months and 6 months.
Funding source	Not specified.
Notes	BASELINE INPUT: (1) CHWs were provided with registration books, standard Ministry of Health of Ethiopia monthly report forms, and stationery sufficient to last a minimum of one year. (2) General meetings were held in each peasant association with the community leaders and the community at large to discuss issues related to the CHW's job description, roles, and responsibilities. (3) Communities were encouraged and helped to establish or strengthen health committees in their respective villages.

Risk of bias

Item	**Reviewers' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "districts were randomly assigned" (p. 380) Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra supervision and refresher training).
Blinding of outcome assessment (detection bias)	Low risk	Quote: "These supervisors were aware of the study but not the district comparison or study hypothesis."
Incomplete outcome data (attrition bias)	Low risk	Comment: 0% CHW attrition (personal communication with author)
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Recruitment bias	Low risk	Comment: No individuals recruited to the health zones post randomisation.
Baseline imbalance	High risk	Comment: baseline comparability of clusters was reported, not adjusted for.
Loss of clusters	Low risk	Comment: no clusters were lost
Unit of analysis issues	High risk	Clustering not taken into account in the analysis

Bailey 1996

Trial reg.	None

For peer review only

Country	Guatemala
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: Promotore
Population	LOCAL TERM FOR CHW: Promotores (Village health promoters) CHWs: N.S. PATIENTS: N.S. HEALTHCARE SETTING: Community INTERVENTION SAMPLE SIZE: 22 CHWs CONTROL SAMPLE SIZE: 27 CHWs
Interventions	OBJECTIVE: Assess the effectiveness of a newly developed method for rapid tiered training of health workers in improving community health worker knowledge and case management skills. INTERVENTION DESCRIPTION: Train CHWs using the interactive "Kader" method, developed in West Java, Indonesia. The Kader method employs tiered training whereby trainers teach trainees exactly as the trainers themselves have been taught. The training relies on role-playing by trainees using "Counselling Cards" to practice counselling or teaching, precisely as the trained village health promoters will counsel mothers in the community. The counselling cards include a diagnostic algorithm and four treatment cards outlining recommendations for five type of diarrhoea. TRAINING DURATION (DAYS): 1 TRAINING DESCRIPTION: Promotores were trained by rural health workers ("Tecnicos") in either interactively or didactically depending on trial arm REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Promotores are supervised by technicos INCENTIVE: None INCENTIVE DESCRIPTION: Promotores are volunteers RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: Train CHWs using traditional didactic training
Outcomes	PATIENTS: None. CHWs: Knowledge score. (The tests consisted of five case descriptions which were used to assess respondents' abilities to identify five diarrhoea types and the appropriate referral and treatment recommendations for each type. General knowledge regarding signs of dehydration, diarrhoea prevention, and methods for making oral rehydration solution was assessed through independent questions.) DURATION OF FOLLOW UP: Immediate
Funding source	Birmingham, Alabama chapter of The Partners of the Americas, the Cadeceus Club of the University of Alabama School of Medicine and the Larry Mayes Memorial Scholarship
Notes	BASELINE INPUT: The responsibilities of the tecnicos include: overseeing vaccination campaigns, education and supervision of promotores, and coordination of health education, promotion, sanitation and water purification efforts in their districts. The promotores are responsible for promoting Department of Public Health programs at the village level and for providing a health advice and referral network for the district health clinics.

Risk of bias

Item	**Reviewers' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "promotores were randomly assigned to receive Kader (experimental) or didactic (control) training" Comment: Unclear how this was done

For peer review only

Allocation concealment (selection bias)	Unclear risk	Comment: Not specified
Blinding of participants and personnel (performance bias)	High risk	Quote: "There is also possible bias secondary to the lack of blinding of participants and trainers"
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Not specified
Incomplete outcome data (attrition bias)	Low risk	Quote: "three promotores had to be excluded from analysis due to incomplete data." Comment: <6% attrition
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol was identified for this review

Bossuroy 2016

Trial reg.	None
Country	India
Methods	STUDY DESIGN: C-RCT (4 arm) UNIT OF RANDOMISATION: Health worker CLUSTERS: 78 (Intervention/Control breakdown not specified)
Population	LOCAL TERM FOR CHW: ASHA CHWs: 72% male. 50% have some university education. Recruited by Delhi-based NGO Operation ASHA PATIENTS: 57.7% male, 57.6% no school or some primary HEALTHCARE SETTING: Community/Operation ASHA centre SAMPLE SIZE: 78 (Intervention/Control breakdown not specified)
Interventions	OBJECTIVE: Investigates whether financial incentives provided to health workers may encourage them to detect new tuberculosis cases and improve treatment adherence. INTERVENTION DESCRIPTION: For the first six months, CHWs were randomly assigned to receive either a fixed salary or a salary dependent on the number of patients they had detected. In the following six months, they were randomly re-assigned to either a fixed or an incentivized salary scheme, based on the number of defaults. CHWs were randomly assigned to one of four treatment arms: they either received (A) financial incentives based on patient detection for six months and incentives based on treatment adherence subsequently, (B) financial incentives based on patient detection for six months and a fixed salary subsequently or (C) a fixed salary for six months and incentives based on treatment adherence subsequently TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: Part of intervention being tested, see above. INCENTIVE DESCRIPTION: Part of intervention being tested, see above. RATIO CHW TO POPULATION: 1:33 START DATE: December 2009 END DATE: April 2013 CONTROL DESCRIPTION: A fixed salary for the whole duration of the experiment.
Outcomes	PATIENTS: TB defaults (number of patients leaving the DOTS system during the

For peer review only

	course of their treatment) CHWs: Case detections (number of patients enrolled in the DOTS system), Health workers' effort/motivation DURATION OF FOLLOW UP: 6 months for each outcome
Funding source	3ie's donors, which include UK aid, the Bill & Melinda Gates Foundation and the Hewlett Foundation.
Notes	Note that the number of months during which each phase was carried out differs across cities, and may not be 6 months in all cases. Also during the second phase, some new CHWs who were not a part of the first phase were randomized into the experiment. BASELINE INPUT: Operation ASHA hires community-based health workers, who are each responsible for operating two DOTS centres. During an initial period of about 3 months, CHWs work on setting up the centres, getting to know the community and surroundings, and making the centre known to the local population. During the first 3 months, CHWs all receive a fixed salary. The experiment starts after the initial three months of a centre lifespan.

Risk of bias

Item	**Reviewers' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "CHWs are randomly assigned" Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear risk	Comment: Unclear how this was done
Blinding of participants and personnel (performance bias)	High risk	Comment: not possible to blind receiving incentives
Blinding of outcome assessment (detection bias)	Low risk	Comment: Two primary outcomes (detection and default) were assessed using TB registers and lab registers kept by public health TB officers
Incomplete outcome data (attrition bias)	Low risk	Quote: "Even though 78 CHWs were initially randomized, 3 CHWs left Op ASHA before their baseline survey…we use all 78 CHWs for impact on attrition"
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol identified
Recruitment bias	Low risk	Comment: CHWs were working before randomisation
Baseline imbalance	Low risk	Comment: Only 5 differences out of 44 are significant at the 10 per cent level and 2 are significant at the 5 per cent level, in line with what should be expected.
Loss of clusters	Unclear risk	Comment: Not adequately described
Unit of analysis issues	Low risk	Quote: "In…all…regressions, we adjust standard errors for clustering at the health worker level since the randomization was conducted at this level."

Carrón 1994

Trial reg.	None
Country	Paraguay
Methods	STUDY DESIGN: C-RCT (2x2 Factorial) UNIT OF RANDOMISATION: Groups (each encompassing 3 districts) CLUSTERS: (A) 3 districts, (B) 3 districts, (C), 3 districts, (D) 3 districts
Population	LOCAL TERM FOR CHW: Community-based distribution (CBD) providers

For peer review only

	CHWs: 100% female. Mean age and mean education level not reported. PATIENTS: Sex, mean age and mean education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: I_a=45 CHWs, I_b=45 CHWs, I_{a+b}=45 CHWs, number of patients unspecified CONTROL SAMPLE SIZE: C=45 CHWs, number of patients unspecified
Interventions	OBJECTIVE: Develop tools of low-cost use to improve the quality of care provided by rural CHWs in Paraguay. Specifically, to determine the relative effects of (1) an interactive counselling guide based on a service algorithm and (2) a supervisory behavioural feedback instrument on the quality of CHW provided care. INTERVENTION DESCRIPTION: Group A (N = 45 CHWs) received (1) Training: use of interactive service algorithm [ABC of the Distributor]. Group B (N = 45 CHWs) received (2) Supervision: use of role-play based instrument to diagnose and provide feedback on CHW service behaviour [Behavioural Feedback Instrument]. Group C (N = 45 CHWs) received (3) Training + Supervision: Use of ABC of the Distributor service algorithm and Behavioural Feedback Instrument-based supervision. INTERVENTION TYPE: Training-Supervision TRAINING DURATION (DAYS): <1 TRAINING DESCRIPTION: (1) ABC of the Distributor: The tool was taught by nurse midwives to CHWs in 30 min individual meetings. (2) Behavioural Feedback Instrument (BFI): CHW were trained by nurse midwives in four 30 min individual meetings. REFRESHER OR ONGOING TRAINING: A two-day course in family planning service provision is imparted once a year, serving as a basic course for untrained distributors and as a refresher course for trained ones. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 07/1992 END DATE: 07/1993 KEY CHW COMPETENCY: CONTROL DESCRIPTION: Group D (N = 45 CHWs) CHWs received the usual two-day annual family planning course and yearly refresher training.
Outcomes	PATIENTS: None. CHWs: Pretest-posttest gains in quality of care provided. Assessed using (1) an observational service test in which a trained simulated client asks services of a CHW and enacts a learnt client profile as she interacts with the provider. When the consultation ends, the simulated client fills out a checklist (range: 41-53 items, depending on client profile) indicating whether the promoter emitted each of the expected provider behaviours; the quality of care score is the sum of the items checked. The service test was administered to all CHWs. (2) A client questionnaire examining 11 indicators across 5 domains: (i) Method choice (user chooses), (ii) Information given to the client (user knowledge of other methods, user knowledge of action upon forgetting, provider advised side effects/method shift), (iii) Interpersonal relations (positive user perceptions), (iv) Provider Technical Competence (physical complaints, contraindications, problems with method, satisfaction with method, unwanted pregnancies), (v) Follow-up mechanisms (provider focus on revisit). The client questionnaire was administered to one patient per CHW randomly selected from the client register. Both the service test and the client questionnaire were conducted by the same trained simulated client. DURATION OF FOLLOW UP: All outcomes collected and reported at baseline and after 12 months.

For peer review only

Funding source	US Agency for International Development (USAID), Office of Population

Risk of bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: groups were "randomly assigned" (p. 10) Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra training or supervision).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: insufficient information to permit judgement
Incomplete outcome data (attrition bias)	High risk	Comment: 156/180 CBD found, 127/180 clients found. Quote: "Provider mortality affected unevenly the distribution of the six client profiles of the Service Test across provider groups. To avoid possible distortions in group comparisons, the equal distribution of client profiles across groups was re-established through a process whereby, for each client profile, (a) the lowest number of cases was determined and (b) the number of cases in the other groups was reduced to this level through random elimination of cases. 124 cases, 31 per group remained in the analysis" (p. 13) Comment: Randomly deleting data is not the same as never having collected the data in the first place; CHWs missing from one group as a result of death may be systematically different than the ones who were randomly deleted from the other groups.
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Recruitment bias	Low risk	Comment: No individuals recruited to groups post randomisation.
Baseline imbalance	Low risk	Comment: Baseline comparability of clusters not reported, pair-matched randomisation of clusters used.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	High risk	Comment: Clustering not taken into account in the analysis.

Chang 2011

Trial reg.	NCT00675389
Country	Uganda
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Clinic catchment area CLUSTERS:
Population	LOCAL TERM FOR CHW: Community-based peer health workers (PHW) CHWs: Sex, mean age, and mean education level not reported. PATIENTS: HIV+ patients 66% female, 34% male. Median age 34 years (range: 15-67) in intervention group and 36 (16-76) in control group.

For peer review only

	Mean education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 13 CHWs, 446 patients CONTROL SAMPLE SIZE: 16 CHWs, 524 patients
Interventions	OBJECTIVE: Determine the effect of an mHealth (mobile phone) CHW support intervention on AIDS care in rural Uganda. INTERVENTION DESCRIPTION: CHWs randomized to the intervention were each given a mobile phone, a one-day residential training, and an hour-long field-based practicum training on the mHealth intervention. These CHWs were asked to send a text message reporting adherence and clinical data back to a centralized database immediately after or during each home visit. These texts were numeric codes produced by converting all quantitative data on the home visit form to simple integers arranged in a predetermined order separated by asterisks, (e.g. pills given, taken, defaulted might convert to *282828*). CHWs in the intervention arm were also encouraged to call a Rakai Health Sciences Program mobile phone hotline or toll-free warmline (staffed only during clinic hours) with questions or concerns. Clinic staff receiving CHW texts and calls could opt to provide care instructions to CHWs, send a higher-level care provider to the patient, or arrange transport to health care facilities. TRAINING DURATION (DAYS): 2 TRAINING DESCRIPTION: PHWs received a residential training on basic HIV pathogenesis, prevention, treatment, adherence counselling, performing pill counts, protecting patient confidentiality, and filling out a home visit form. In addition to their baseline training, PHWs randomised to the mHealth Arm were each given a mobile phone, a one day residential training, and an hour long field-based practicum training on the mHealth intervention. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: PHW day-to-day supervision was largely performed by a single staff member working part-time. INCENTIVE: Financial INCENTIVE DESCRIPTION: Monthly allowance of about 12.5 USD RATIO CHW TO POPULATION: 1:34 START DATE: 04/2006 END DATE: 06/2008 CONTROL DESCRIPTION: CHWs who did not receive the mHealth intervention continued providing regular care (see baseline input).
Outcomes	PATIENTS: Primary: cumulative risk of virologic failure (any failure during follow-up period equalling failure with viral loads conducted every 24 weeks; failure defined as .400 copies/mL). Secondary: (i) patient adherence (pill counts and 3 day self-report), (ii) virologic failure at 24 and 48 weeks of ART (failure defined as 400 copies/ml), (iii) lost to follow- up (lack of a pharmacy refill visit in over 90 days), and (iv) mortality (ascertained through verbal autopsies) DURATION OF FOLLOW UP: A summary clinic pill count was calculated by dividing the number of pills taken over the study period by the sum of pills expected to be taken over the study period. Viral loads were measured at 24 to 192 week time points from antiretroviral therapy initiation by 24-week intervals. CHWs: None.
Funding source	Doris Duke Charitable Foundation, The Division of Intramural Research, The National Institute for Allergy and Infectious Diseases, National

For peer review only

	Institutes of Health, and National Institutes of Health Training (2T32-AI07291) and Career Development (1K23MH086338-01A2) Grants.
Notes	BASELINE INPUT: CHWs received a two-day residential training on basic HIV pathogenesis, prevention, treatment, adherence counselling, performing pill counts, protecting patient confidentiality, and filling out a home visit form. CHWs provided clinical and adherence monitoring and psychosocial support to fellow patients. At home visits, CHWs recorded a review of symptoms, an adherence self-report, and pill count results using a standardized form. CHWs were each given a bicycle, t-shirts, basic supplies, and an initial monthly allowance of about 12.5 USD. CHW day-to-day supervision was largely performed by a single RHSP staff member working part-time.

Risk of bias

Item	**Reviewers' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "29 PHWs at 10 clinics were randomized by clinic to receive the intervention or not." Comment: Unclear how done.
Allocation concealment (selection bias)	Low risk	Quote (parent trial): We assigned clusters using unmatched, unrestricted random allocation by a drawing of lots. Comment: Author provided the following information via email: "allocation was done simultaneously with the randomisation process"
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra training and a mobile phone).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: Author provided the following information via email: "this was a biological sample…no special measures taken"
Incomplete outcome data (attrition bias)	Low risk	Quote (parent trial): 9.3% died; 2.2 % LTFU. "Efficacy analyses for cumulative risk of virologic failure and for virologic outcomes at specific time points from ART initiation were by intention to treat using log-binomial regression with generalized estimating equations (GEE)"
Selective reporting (reporting bias)	Low risk	Comment: Parent trial registered on ClinicalTrials.gov (NCT00675389). All outcomes reported.
Recruitment bias	Low risk	Comment: No individuals recruited to clinic catchment area post randomisation.
Baseline imbalance	Low risk	Quote: Sociodemographic characteristics, immunologic and clinical stage of disease, the proportion of patients on ART, the median duration of time patients were on ART prior to the start of the trial, and the pre-trial 24 week and 48 week virologic failure" were well balanced between arms.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	Low risk	Quote: "…robust variance estimates appropriate for cluster-randomised trials"

For peer review only

DeRenzi 2012A

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: RCT (Pilot study) UNIT OF RANDOMISATION: CHW
Participants	LOCAL TERM FOR CHW: community health worker CHWs: 100% Female. Age range 23-55 years. Mean schooling unreported. PATIENTS: Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 8 CHWs, number of patients unspecified CONTROL SAMPLE SIZE: 7 CHWs, number of patients unspecified
Interventions	OBJECTIVE: Evaluate the impact of SMS reminders to improve the promptness of routine CHW visits. INTERVENTION DESCRIPTION: CHWs received an SMS message two days after they had reported a referral with CommCare, reminding them of the follow-up visit. Daily SMS messages were sent until the follow-up was recorded. INTERVENTION TYPE: Supervision TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: Non-financial INCENTIVE DESCRIPTION: CHWs are incentivized with personal use of the mobile phone they are given for CommCare. RATIO CHW TO POPULATION: N.S. START DATE: 04/2010 END DATE: 06/2010 CONTROL DESCRIPTION: CHWs received nothing.
Outcomes	PATIENTS: Change in percentage of closed referrals CHWs: None TIMEPOINTS COLLECTED/REPORTED: Referral and follow-up information was collected automatically by CHWs using CommCare, and sent immediately to the server by the CHW. 24 days
Funding source	Not specified.
Notes	The study period was split into baseline and intervention periods that were 39 and 24 days long, respectively. Two CHWs in the control group were excluded from analysis because they did not report any referrals during the intervention period.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: " participants were randomly assigned" (p. 27) Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra SMS notifications)
Blinding of outcome assessment (detection	Unclear risk	Comment: insufficient information to permit judgement.

For peer review only

bias)		
Incomplete outcome data (attrition bias)	Low risk	Comment: "Two CHWs in the control group were excluded from analysis because they did not report any referrals during the intervention period."
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement

DeRenzi 2012B

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: CHW (stratified by baseline performance)
Participants	LOCAL TERM FOR CHW: community health worker CHWs: Male and Female (proportions unspecified. Age range 23-63 years. Mean schooling unreported. PATIENTS: Chronic care patients (most are HIV-positive, though some have diabetes, tuberculosis, or other long-term and chronic ailments). Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION AND CONTROL SAMPLE SIZE: 87 CHWs (breakdown unspecified, after attrition: I =34, C = 40), number of patients unspecified
Interventions	OBJECTIVE: Evaluate the impact of SMS reminders to improve the promptness of routine CHW visits. INTERVENTION DESCRIPTION: Escalating reminder system that sent SMS reminders directly to the CHW before notifying the CHW's supervisor after several overdue days (augments existing supervision structures). INTERVENTION TYPE: Supervision TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: Non-financial INCENTIVE DESCRIPTION: CHWs are incentivized with personal use of the mobile phone they are given for CommCare. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: CHWs received nothing after baseline input.
Outcomes	PATIENTS: Mean time chronic care patients overdue. CHWs: None. TIMEPOINTS COLLECTED/REPORTED: Data collected on a rolling basis, sent from CHWs post closed referral (40 days)
Funding source	Not specified.
Notes	The study period was split into baseline and intervention periods that were 40 days each. BASELINE INPUT: Pathfinder CHWs are all volunteers; they are provided with a "generous transportation stipend" during their monthly meetings and incentivized with personal use of the mobile phone they are given for CommCare. CHANGES FROM PILOT: In the pilot study (DeRenzi 2012A), the earliest reminder was sent on the evening that the CHW's visit was due, too late to make an on-time visit. For the Dar es Salaam studies (DeRenzi 2012B, DeRenzi 2012C), *proactive* reminders were incorporated, with the first one being sent the day before a visit is due. Second, if a referral was not closed during the pilot, the system continued to send SMS messages but had no means of determining *why*

For peer review only

	the CHW was not reporting a follow-up. This issue directly inspired the escalation to supervisor evaluated in DeRenzi 2012B.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: " CHWs in each bin were then randomly assigned to the intervention or control group." Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra SMS messages and supervisor intervention).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: insufficient information to permit judgement.
Incomplete outcome data (attrition bias)	High risk	Quote: "Although 87 CHWs participated in this study, 13 participants were removed before analysis for one of the following reasons: no data submitted during the intervention period (11 CHWs), no active cases (1 CHW), and technical issues with CommCare or the phone (1 CHW). This removal left 34 CHWs in the SMS+Supervisor group and 40 CHWs in the control group." "During the course of the intervention, there were occasional phone problems (e.g. accidentally deleting the CommCare application) and discrepancies (e.g. a CHW reporting s/he had already sent data) reported from CHWs in the field. If there was a discrepancy, we removed the client from our analysis. This happened for 27 clients during the intervention period, which represented 3.7% of the total number of clients visited during that period."
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement

DeRenzi 2012C

Trial reg.	None
Country	Tanzania
Methods	STUDY DESIGN: RCT UNIT OF ALLOCATION: CHW
Participants	LOCAL TERM FOR CHW: Community health worker CHWs: Male and Female (proportions unspecified. Age range 23-63 years. Mean schooling unreported. PATIENTS: Chronic care patients (most are HIV-positive, though some have diabetes, tuberculosis, or other long-term and chronic ailments). Sex, age and education level not reported. HEALTHCARE SETTING: Household INTERVENTION AND CONTROL SAMPLE SIZE: 87 CHWs (breakdown unspecified), number of patients unspecified
Interventions	OBJECTVE: Although only 6.5% of cases in DeRenzi 2012B required the involvement of the supervisor, this escalation step is still more costly and burdensome than sending SMS messages alone. The objective of this study was to

For peer review only

	quantify the effect of escalating to the supervisor versus SMS messages alone in the context of SMS reminders to improve the promptness of routine CHW visits. INTERVENTION: Intervention group (SMS-Only) was provided only with proactive and overdue reminders, but escalation to the supervisor never occurred. MODE OF DELIVERY: The system, implemented on top of CommCare, first sends proactive reminders to intervention CHWs (the day before (x1) and the day of (x2) a scheduled routine visit. Daily reminders are sent while the visit remains overdue, informing the CHW od the # of days the visit is overdue. INTERVENTION TYPE: Supervision TRAINING DURATION (DAYS): N.S. TRAINING DESCRIPTION: N.S. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: Part of intervention SUPERVISION DESCRIPTION: Part of intervention INCENTIVE: Non-financial INCENTIVE DESCRIPTION: CHWs are incentivized with personal use of the mobile phone they are given for CommCare. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL: CHWs (SMS+ Supervisor) received exactly the same input as the intervention group in DeRenzi 2012B.
Outcomes	PATIENTS: None. CHWs: Mean time overdue. TIMEPOINTS COLLECTED/REPORTED: 90 days
Funding source	Not specified.
Notes	90-day baseline period and a 90-day intervention period. BASELINE INPUT: Since the baseline period followed directly after DeRenzi 2012B and with the same CHWs, some CHWs continued to receive escalating SMS reminders during the baseline, while others did not. H1: For CHWs who received escalating SMS reminders during the baseline period, the SMS-Only intervention (removing the supervisor) would decrease performance compared to continuing in the SMS+Supervisor condition. H2: For CHWs who did *not* any receive SMS reminders during the baseline period; the SMS+Supervisor intervention would result in a larger increase in performance than the SMS-Only intervention. The studies presented here measure reported follow-ups and are not correlated with ground truth of actual follow-ups.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: " CHWs were then randomly assigned to one of two conditions" Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received SMS and/or additional supervision).
Blinding of outcome assessment (detection bias)	Unclear risk	Comment: insufficient information to permit judgement.

For peer review only

Incomplete outcome data (attrition bias)	High risk	Quote: "Of the original 87 CHWs, 26 were excluded from analysis for the following reasons: incorrect phone number (2 CHWs), having no data during baseline, intervention, or both periods (22 CHWs), and technical problems with the phone (2 CHWs). In the end, there were 32 CHWs originally from the intervention group in Study 1 (16 SMS+Supervisor, 16 SMS-Only) and 29 from the control group in Study 1 (14 SMS+Supervisor, 15 SMS-Only)."
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement

Gautham 2015

Trial reg.	N.S.
Country	India
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: CHW
Population	LOCAL TERM FOR CHW: Rural health provider CHWS: 8 male, 8 female; existing rural health providers; Males: mean age of 48.75 years, 10 years of education or more. Females: mean age 24.75 years, 10 years of education or more PATIENTS: Paediatric and Adult patients HEALTHCARE SETTING: Community (personal communication with S. Iyengar) INTERVENTION SAMPLE SIZE: 8 CONTROL SAMPLE SIZE: 8
Intervention	OBJECTIVE: (a) Use a randomized control design to measure changes in protocol compliance (PC) by health workers in their everyday work settings, (b) assess the usability and acceptability of the mobile application with health workers in the experimental group and (c) obtain patient feedback on health workers' use of the mobile system during treatment. INTERVENTION DESCRIPTION: Mobile phone–based, media-rich procedural guidance applications TRAINING DURATION (DAYS): 2 days TRAINING DESCRIPTION: Training programme in guideline-based care and in use of the mHealth system by physicians who spoke the local language. The study team assessed all RHPs on their understanding of guideline-based care using a simple assessment instrument developed by the research team, on which every RHP received a pre-and post-training test score (PTTS). REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: N.S. END DATE: N.S. CONTROL DESCRIPTION: A set of paper guidelines to use in the field
Outcomes	PATIENTS: None

For peer review only

	CHWs: Protocol compliance DURATION OF FOLLOW UP: 2 months
Funding	Information Society Innovation Fund (ISIF)

Risk of Bias

Item	Author's judgement	Description
Random sequence generation (selection bias)	Unclear risk	Quote: "At the end of the training, we randomly assigned half the RHPs to the control group and other half to an experimental group." Comment: Unclear how this was done
Allocation concealment (selection bias)	Unclear risk	Quote: "At the end of the training, we randomly assigned half the RHPs to the control group and other half to an experimental group." Comment: Unclear how this was done
Blinding of participants and personnel (performance bias)	High risk	Quote: "after the training, the RHPs in the experimental group went back to their field sites with the applications on their mobile phones, while for the ones in the control group, the data card with the application was taken out" Comment: Unclear if patients were blinded. CHWs were not blinded.
Blinding of outcome assessment (detection bias)	High risk	Quote: "PC was evaluated by direct observations of the RHPs' performance as they were treating the patients."
Incomplete outcome data (attrition bias)	Unclear risk	Comment: Not described
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol identified

Ginsburg 2014

Trial reg.	N.S.
Country	Bangladesh
Methods	STUDY DESIGN: RCT UNIT OF RANDOMISATION: CHW
Population	LOCAL TERM FOR CHW: Community health workers CHWS: Female; age N.S.; high school education; recruited from rural villages by study team PATIENTS: Adult women aged 25 and older and living in the catchment area of Khulna Division, Bangladesh. Exclusion criteria were women younger than age 25 and women unable to give verbal consent. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: A: 246 B: 240 CONTROL SAMPLE SIZE: C: 70
Intervention	OBJECTIVE: Test whether CHWs guided by smart phone applications would be more efficient and effective than CHWs without phones at encouraging women with an abnormal breast exam to adhere to advice regarding a clinic appointment INTERVENTION DESCRIPTION: Arm A: smart phone with applications to guide interview, report data, show motivational video, and offer an appointment for women with an abnormal CBE. Arm B: smart phone platform as in arm A plus additional "patient navigation training."

For peer review only

	TRAINING DURATION (DAYS): 3 TRAINING DESCRIPTION: CHWs and supervisors attended an intensive training program on breast anatomy, physiology, and pathology, as well as the 2-hour training session on CBEs. CHWs were taught in a gentle, culturally sensitive manner how to obtain informed verbal consent, to ask the structured interview questions, and to perform a CBE. Only the 10 CHWs randomized to arm B were given an extra day of training in patient navigation. REFRESHER OR ONGOING TRAINING: N.S. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Three female CHW supervisors, also local women but with postsecondary education and prior experience in the health or social services sector, were hired to oversee daily workflow and to liaise with our field operations team. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: 1:19 START DATE: July, 2012 END DATE: October, 2012 CONTROL DESCRIPTION: No smart phone; same interview recorded on paper
Outcomes	PATIENTS: The "adherence" (to advice regarding a clinic appointment) for women with an abnormal CBE CHWs: N.S. DURATION OF FOLLOW UP: 1 month
Funding	Rising Stars in Global Health award from Grand Challenges Canada.

Risk of Bias

Item	Author's judgement	Description
Random sequence generation (selection bias)	Low Risk	Quote: "The 30 CHWs were randomized by manual lottery (by O.M.G., eyes closed, hand in jar with identical folded papers with CHW names) to one of the three study arms."
Allocation concealment (selection bias)	Low Risk	Comment: Simultaneous with randomisation
Blinding of participants and personnel (performance bias)	Unclear	Comment: Unclear if patients were blinded. CHWs could not be blinded.
Blinding of outcome assessment (detection bias)	Low risk	Comment: clinic records used, so low risk of bias
Incomplete outcome data (attrition bias)	Unclear risk	Comment: No attrition reported
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol identified

Omer 1998

Trial reg.	None
Country	Pakistan
Methods	STUDY DESIGN: C-RCT UNIT OF RANDOMISATION: Enumeration area CLUSTERS: Intervention: 5, Control: 5
Population	LOCAL TERM FOR CHW: Lady health workers (LHWs)

For peer review only

	CHWs: 100% female. Mean age and education level not reported. PATIENTS: 100% women, of whom 45% report receiving "some formal education." Mean age not reported. HEALTHCARE SETTING: Household INTERVENTION SAMPLE SIZE: 52 CHWs, 529 patients CONTROL SAMPLE SIZE: Unspecified number of CHWs, 541 patients
Interventions	OBJECTIVE: Demonstrate the effective use of community-based evidence for health promotion by CHWs in Sindh, Pakistan. INTERVENTION DESCRIPTION: In a consultation forum convened for the purpose, CHWs and government health officials developed a targeted tool to guide household discussions. It was designed to be "user-friendly, indigenous, and attractive" to both CHWs and their clients. The tool was based on ajrak, a traditional Sindh fabric, with the message in embroidery. The design team also consulted with the provincial team of the National Programme, CHWs, women in different urban and rural communities, local designers and handicraft experts. The embroidery depicted maternal practices like attending and not attending antenatal check-ups, giving colostrum after birth and not doing heavy work. The team pre-tested the embroidery in five communities not included in this study, ensuring input from CHWs and women in the communities to adjust the tools. In intervention communities, CHWs showed these panels to pregnant women during their household visits and invited them to describe what was going on in the pictures. They encouraged their clients to express their views of the situations depicted on the cloths and discussed any difficulties they might face. To go for prenatal check-ups, for example, the discussion might include when and how the person would go. Training in the tool included an explanation of the tools, role modelling and practice in the community. INTERVENTION TYPE: Training/integration TRAINING DURATION (DAYS): Part of intervention TRAINING DESCRIPTION: Part of the intervention. REFRESHER OR ONGOING TRAINING: Part of the intervention. SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: Monitoring relied on the existing LHW supervisory structure in both intervention and control communities. As part of routine control visits, supervisors did random checks to see if LHWs used the tools in intervention communities. LHSs held monthly meetings at district level, where they would get feedback on progress and problems they were not able to rectify on their own. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: N.S. START DATE: 06/2000 END DATE: 04/2001 CONTROL DESCRIPTION: CHWs (N = unspecified) in control communities did not use the embroidered cloths and received no training in how to use them, but continued their usual work visiting women before and after childbirth to discuss safe practices in pregnancy based on the standard training of the CHW programme. (N = 541 patients)
Outcomes	PATIENTS: (i) attendance at prenatal checkups; (ii) stopping of routine heavy work during pregnancy (iii) giving colostrum to newborn babies (iv) maintaining exclusive breastfeeding for four months. CHWs: Performance (# of visits to pregnant women) DURATION OF FOLLOW UP: All outcomes were collected and reported at baseline and in a household survey after CHWs in intervention communities had used the tools for 10 months.
Funding source	Canadian International Development Agency's (CIDA) Canada Fund for Local

For peer review only

	Initiatives in Pakistan.

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "We allocated these randomly using a computerised random number generator" (p. 180)
Allocation concealment (selection bias)	Unclear risk	Quote: "...using a computerised random number generator" (p. 180) Comment: Unclear if enumeration areas were assessed before number appeared.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra training on a novel communication tool).
Blinding of outcome assessment (detection bias)	Low risk	"Fieldworkers, who did not know which communities were intervention and which were control, collected data" (p. 180)
Incomplete outcome data (attrition bias)	Low risk	Quote: "Analysis followed the intention to treat principle, with all women in the intervention communities analysed as though exposed, whether or not they had seen the intervention materials." (p. 180)
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Other bias	Low risk	In cluster-randomized trials, particular biases to consider include: (i) Recruitment bias Comment: Low risk; no individuals recruited to enumeration area post randomisation. (ii) Baseline imbalance Comment: Unclear risk; baseline comparability of clusters not reported, not adjusted for. (iii) Loss of clusters Comment: Low risk; no clusters were lost (iv) Unit of analysis issues Comment: Unclear risk; not clear if clustering taken into account in the analysis. (v) Comparability with individually randomized trials Comment: Irrelevant, all studies in this review are C-RCTs.
Recruitment bias	Low risk	Comment: No individuals recruited to enumeration area post randomisation.
Baseline imbalance	Unclear risk	Comment: Baseline comparability of clusters not reported, not adjusted for.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	Unclear risk	Comment: Not clear if clustering taken into account in the analysis.

Winch 2003

Trial reg.	None
Country	Mali
Methods	STUDY DESIGN: Pair matched C-RCT UNIT OF RANDOMISATION: Health zones CLUSTERS: Intervention: 5; Control: 5

For peer review only

	DURATION: 15 months (Visual aids: 11/2001-02/2002; Referral mechanism: 03/2001-06/2002)
Population	LOCAL TERM FOR CHW: Gérant (village drug kit managers) CHWs: Sex, age and education level not reported. PATIENTS: The intervention was directed at children <14 years. Mean schooling of carers was 0.85 years. Sex of all patients treated was not reported. HEALTHCARE SETTING: Household INTERVENTION AND CONTROL SAMPLE SIZE: ~102 CHWs in both arms, 152 patients in the intervention arm, 134 in the control arm
Interventions	OBJECTIVE: Evaluate an intervention to (i) improve the skills of CHWs to counsel parents on correct home administration of chloroquine (CQ), and (ii) increase the referral of sick children to community health centres (CHC). INTERVENTION DESCRIPTION: (1) Use of visual aids for drug administration and danger sign recognition (2) Introduction of a referral/counter referral recordkeeping mechanism (3) Refresher training (4) Community engagement INTERVENTION TYPE: Training-Reporting- Integration TRAINING DURATION (DAYS): Part of the intervention TRAINING DESCRIPTION: Part of the intervention REFRESHER OR ONGOING TRAINING: a 5 day additional training course was conducted in a local elementary school that covered: (i) doses of CQ for different age groups (also covered in the standard course), (ii) how to counsel carers of young children regarding the correct administration of CQ in the home and the importance of going to the CHC when referred, (iii) indications for referral including danger signs and failure of the child's symptoms to respond to treatment, (iv) basic evaluation of a sick child for danger signs such as dehydration requiring referral, and (v) how to use the referral notebook and the visual aids SUPERVISION FREQUENCY: N.S. SUPERVISION DESCRIPTION: N.S. INCENTIVE: N.S. INCENTIVE DESCRIPTION: N.S. RATIO CHW TO POPULATION: ~1:3 START DATE: 03/2001 END DATE: 06/2002 CONTROL DESCRIPTION: When a new village drug kit is established, all managers go through standard training which consists of a 35-d literacy course in which they are taught to read and write the national language, Bambara, using a phonetic alphabet, followed by a 1-week course on how to manage the drug kit including the correct doses to recommend and sell for each product in the kit. All of the drug kit managers in both the intervention and comparison groups had had this standard training prior to the study.
Outcomes	PATIENTS: The proportion of children with indications for referral to a health facility who were referred to and subsequently reached the facility. CHWs: (i) The proportion of children who were treated with a full course of CQ in the home DURATION OF FOLLOW UP: At least twice a week an interviewer visited the CHW, talked to him/her and examined the register of sales maintained, noting down the cases where CQ had been sold to treat a child aged <6 years. The interviewer assed the effect of the intervention on the quality and content of the counselling received from the CHW through direct questions during follow-up in the home 5 days after the sale. Data from health centre registers and CHW referral notebooks was collected at the end of the trial. All data was reported at last observation.
Funding source	United States Agency for International Development through the Family Health and Child Survival Cooperative Agreement with Johns Hopkins Bloomberg School of Public Health

For peer review only

Risk of Bias

Item	Reviewers' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "1 health zone from each pair was randomly assigned to receive the intervention." Comment: Unclear how done.
Allocation concealment (selection bias)	Unclear risk	Comment: insufficient information to permit judgement, method of concealment is not described.
Blinding of participants and personnel (performance bias)	High risk	Comment: participants and personnel could not be blinded (they received extra materials, refresher training and referral recording materials).
Blinding of outcome assessment (detection bias)	Unclear risk	Quote: "Fifteen field interviewers were recruited. In 6 of the 10 health zones there was 1 interviewer only, while the other 4 health zones had 2 or 3 interviewers due to their large size. Each interviewer was assigned a set of village drug kits and their clients to study. At least twice a week the interviewer visited the manager of the village drug kit, talked to him/her and examined the register of sales maintained by him/her, noting down the cases where CQ had been sold to treat a child aged <6 years. On the fifth day after the sale, the field interviewer went to the child's house and interviewed the person who had administered the treatment to the child. " (p. 483) Comment: Insufficient information to permit judgement of 'Low risk' or 'High risk'
Incomplete outcome data (attrition bias)	Unclear risk	Comment: Insufficient reporting of attrition/exclusions to permit judgement.
Selective reporting (reporting bias)	Unclear risk	Comment: Insufficient information to permit judgement
Recruitment bias	Low risk	Comment: No individuals recruited to health zones post randomisation.
Baseline imbalance	Low risk	Comment: Baseline comparability of clusters not reported, pair-matched randomisation of clusters used.
Loss of clusters	Low risk	Comment: No clusters were lost
Unit of analysis issues	Low risk	Comment: Intracluster correlation coefficient taken into account in the analysis.

Appendix F: Characteristics of Excluded, Ongoing & Studies Awaiting Classification

Excluded studies:

Record name	Year	Reason for exclusion
Afenyadu	2005	Did not meet the definition for CHW (CHWs were teachers)
Ahmed	1993	Not a randomised controlled trial

For peer review only

Akogun	2001	Not a randomised controlled trial
Alam	2012	Not a randomised controlled trial
Amouzou	2014	Non-CHW comparator
Ashraf	2014	Did not meet the definition for CHW (CHWs were hair stylists)
Avula	2011	Not a randomised controlled trial
Baqui	2008	Non CHW comparator
Begum	2009	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Bertrand	1986	Non CHW comparator
Bhatt	1981	Not a randomised controlled trial
Bhutta	2008	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Björkman	2009	Health workers were primarily facility-based
Bojang	2011	Non CHW comparator
Boone	2008	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Brieger	2000	Health workers were primarily facility-based
Chin-Quee	2016	Not a randomised controlled trial
Curtale	1995	Not a randomised controlled trial
Das	2014	Effects of the CHW performance intervention cannot be disaggregated
Davies-Adetugbo	1997	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Elder	1992	Not a randomised controlled trial
Findley	2013	Not a randomised controlled trial
Florez-Arango	2011	Did not meet the definition for CHW (one quarter of assessed health workers were primarily facility-based)
Foreit	1984	Health workers were primarily facility-based
George	2012	Did not meet the definition for CHW (Not part of health system)
Harvey	2008	Not a randomised controlled trial
Iyengar	2013	One quarter of assessed health workers were primarily facility-based
Jintaganont	1988	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Joos	2016	Performance interventions in both arms
Jung	2009	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance

For peer review only

Kalyango	2012	Intervention not designed to improve performance
Morris	2015	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Owais	2011	Tested the efficacy of an intervention for mothers rather than efficacy of an intervention to improve CHW performance
Pineda	1983	Not a randomised controlled trial
Pradeep	2014	Non-CHW comparator
Puett	2012	Not a randomised controlled trial
Sranacharoen-pong	2012	Intervention not designed to improve performance
Turab	2014	Tested the efficacy of CHWs for delivering a health intervention rather than efficacy of an intervention to improve CHW performance
Zafar	2016	Intervention not designed to improve performance
Zurovac	2011	Health workers were primarily facility based

Studies awaiting classification:

Author	Year	Title
Carrón	1993	Desarrollando instrumentos de bajo costo para mejorar la calidad de los servicios rurales de planificación familiar: informe final (Enero 1992 - Diciembre 1993).
Uwima	2013	Community-based intervention to enhance provision of integrated TB-HIV and PMTCT services in South Africa.

Ongoing studies:

Hackett 2015

Trial Registration	N.S.
Trial name	Job satisfaction, self-efficacy, and performance of community health workers participating in a mobile health (mHealth) program to improve maternal, newborn and child health (MNCH) in Rural Tanzania
Objective	Evaluate the process and impact of an mHealth intervention
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: CHW pairs
Country	Tanzania
Participants	LOCAL TERM FOR CHW: Community health worker CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: mHealth KEY CHW COMPETENCY: CHWs were trained on a smartphone application designed to improve data

For peer review only

	management, patient tracking, and delivery of key health/nutrition messages to pregnant women and mothers.
Control	Ministry of Health's national community- MNCH program.
Outcomes	PATIENTS: Satisfaction CHWs: Performance
Starting date (dd/mm/yyyy)	N.S.
Contact information	Kristy Hackett PhD Candidate. University of Toronto Toronto, ON. M5T 3M7. Canada. +1-416-978-2058 kristy.hackett@mail.utoronto.ca
Funding Source	Foreign Affairs, Trade and Development Canada (DFATD) W. Garfield Weston Doctoral Fellowship Program

Källander 2012

Trial Registration	NCT01972321
Trial name	Evaluating the effect of innovative motivation and supervision approaches on community health worker performance and retention in Uganda and Mozambique: study protocol for a randomised controlled trial
Objective	Increase sustainable coverage of ICCM in Uganda and Mozambique by designing and evaluating innovations for increased CHW supervision and motivation.
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Sub-counties
Country	Mozambique and Uganda
Population	LOCAL TERM FOR CHW: Community health worker CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: mHealth KEY CHW COMPETENCY: The mHealth interventions in Uganda and Mozambique encompass three main activities: 1) closed user groups to enable free two-way communication between CHWs and their supervisors; 2) weekly ICCM data submission using phones with automated motivational feedback, SMS to supervisors flagging problems for target supervision, and summary ICCM statistics made accessible online to district statisticians; and 3) monthly motivational and constructive SMS to CHWs.
Control	Children living in the control areas continued to benefit from the routine Ministry of Health iCCM package provided by the CHWs who were supported by the national and sub-national health services with funding from Malaria Consortium's CIDA-iCCM project.

For peer review only

Outcomes	PATIENTS: Proportion of sick children appropriately treated. CHWs: (1) Performance, (2) Motivation
Starting date (dd/mm/yyyy)	11/2009
Contact information	Karin Källander, PhD +256772744126 k.kallander@malariaconsortium.org
Funding source	Bill & Melinda Gates Foundation and DFID Programme Partnership Agreement (DFID-PPA)

Rabbani 2015

Trial registration	ACTRN12613001261707
Trial name	Improving community case management of diarrhoea and pneumonia in district Badin, Pakistan through a cluster randomised study
Objective	Evaluate the affect of new supervisory strategies employed by Lady Health Supervisors
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Lady health supervisors
Country	Pakistan
Population	LOCAL TERM FOR CHW: Lady health workers CHWs: Women. Age and education level not reported. PATIENTS: Children. Sex and education level not reported.
Interventions	INTERVENTION DESCRIPTION: The intervention consists of training to build LHS knowledge and skills, clinical mentorship and written feedback to LHWs KEY CHW COMPETENCY: Community case management
Control	Standard care
Outcomes	PATIENTS: None CHWs: Improvement in CCM practices of childhood diarrhoea and pneumonia
Starting date (dd/mm/yyyy)	11/2014
Contact information	Fauziah Rabbani fauziah.rabbani@aku.edu Department of Community Health Sciences, The Aga Khan University, Stadium Road, 3500, Karachi 74800, Pakistan
Funding source	WHO Geneva, Department of Maternal, Newborn, Child and Adolescent Health.

Saha 2013

Trial registration	N.S.
Trial name (registration)	A Comprehensive Training And Supervision Program Improved Frontline Health Workers' Knowledge Of Infant And Young Child Feeding In Rural Bangladesh

For peer review only

Objective	Increase sustainable coverage of ICCM in Uganda and Mozambique by designing and evaluating innovations for increased CHW supervision and motivation.
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Not reported
Country	Bangladesh
Population	LOCAL TERM FOR CHW: Shasthya Sebika & Pushti Kormi CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: Education, Incentives, Supervision KEY CHW COMPETENCY: A comprehensive infant and young child feeding practices training and performance improvement package (3-day classroom discussion, 3-day field practice, monthly refreshers, routine supervision and performance-based incentives) plus a mass media campaign.
Control	Mass media campaign
Outcomes	PATIENTS: None CHWs: CHW knowledge of infant and young child feeding practices
Starting date (dd/mm/yyy)	2010
Contact information	Kuntal Kumar Saha, International Food Policy Research Institute (IFPRI) kuntalkumar38@gmail.com
Funding source	Bill and Melinda Gates Foundation

Vedanthan 2014

Trial registration	NCT01844596
Trial name	Optimizing linkage and retention to hypertension care in rural Kenya (LARK hypertension study): study protocol for a randomized controlled trial
Objective	Investigate whether community health workers, equipped with a tailored behavioural communication strategy and smartphone technology, can increase linkage and retention of hypertensive individuals to a hypertension care program and significantly reduce blood pressure among them.
Methods	STUDY DESIGN: Cluster-randomized trial UNIT OF RANDOMISATION: Community
Country	Mozambique and Uganda
Population	LOCAL TERM FOR CHW: Community health worker CHWs: Sex, age and education level not reported. PATIENTS: Sex, age and education level not reported.
Interventions	INTERVENTION DESCRIPTION: 1) community health workers with an additional tailored behavioural

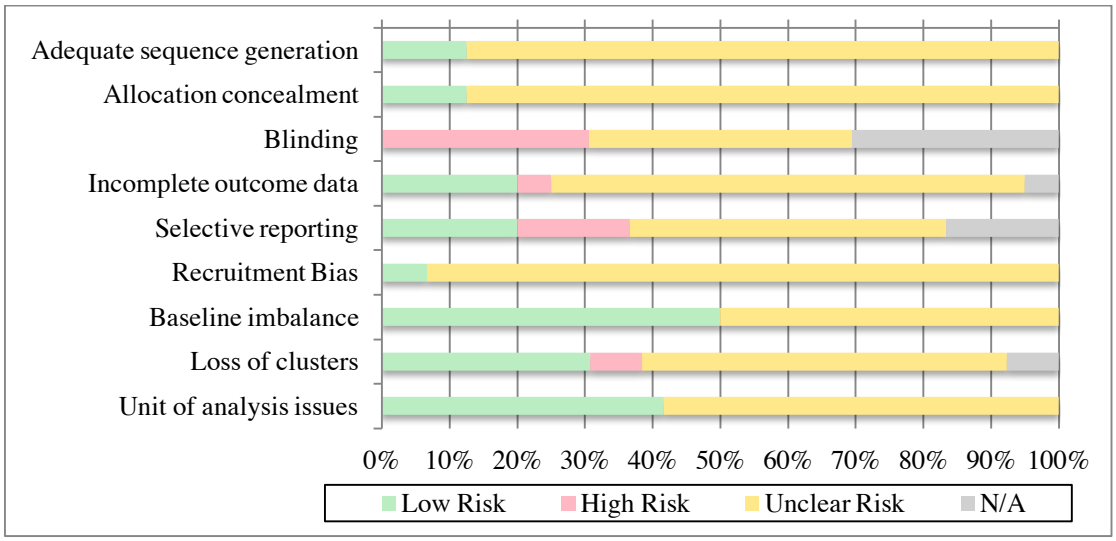
For peer review only

	communication strategy; and 2) community health workers with a tailored behavioural communication strategy who are also equipped with smartphone technology. KEY CHW COMPETENCY: Motivational interviewing and tailored communication- strategies
Control	Usual care (community health workers with the standard level of hypertension care training)
Outcomes	PATIENTS: The co-primary outcome measures are: 1) linkage to hypertension care, and 2) one-year change in systolic blood pressure among hypertensive individuals. Cost-effectiveness analysis will be conducted in terms of costs per unit decrease in blood pressure and costs per disability-adjusted life year gained. CHWs: None.
Starting date (dd/mm/yyyy)	04/2014
Contact information	Rajesh Vedanthan rajesh.vedanthan@mssm.edu Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1030, 10029 New York, USA
Funding source	National Heart, Lung, and Blood Institute of the National Institutes of Health under award number 1U01HL114200, United States, under The Global Alliance for Chronic Diseases programme

Appendix G: Risk of bias of Included Studies

Risk of bias assessments are included in the Characteristics of included studies tables in Appendix E and are summarised in the figure and table below. These assessments were used for interpreting the results and assessing the quality of evidence for specific outcomes. The first six categories apply to all included trials, while the latter four apply only to cluster-randomised control trials.

For peer review only



Review author's judgement about each methodological quality item presented as percentages across all included studies

	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Recruitment bias	Baseline imbalance	Loss of clusters	Unit of analysis issues
Andreoni (2016)	?	?	?	?	+	?				
Ashraf (2015)	?	?	?	?	-	?	+	?	?	+
Ayele (1993)	?	?	-	+	+	?	+	-	+	-
Bailey (1996)	?	?	-	?	+	?				
Bossuroy (2016)	?	?	-	+	+	?	+	+	?	+
Carrón (1994)	?	?	-	?	-	?	+	+	+	-
Chang (2011)	?	+	-	?	+	+	+	+	+	+
DeRenzi (2012A)	?	?	-	?	-	?				
DeRenzi (2012B)	?	?	-	?	-	?				
DeRenzi (2012C)	?	?	-	?	-	?				
Gautham (2015)	?	?	-	-	?	?				
Ginsburg (2014)	+	+	?	+	?	?				
Omer (1998)	+	?	-	+	+	?	+	?	+	?
Winch (2003)	?	?	-	?	?	?	+	+	+	+

For peer review only

Methodological quality summary: review author's judgement about each methodological quality item for each included study

Random sequence generation and allocation concealment (selection bias)

Though all trials reported that they were randomised, only one two described the method of randomisation (Ginsburg et al., 2014; Omer et al., 2008). Likewise, only two trials reported a method of allocation concealment (Chang et al., 2011; Ginsburg et al., 2014). As such, the potential for selection bias is a notable weakness of this data set.

Blinding (performance bias and detection bias)

Given the time and human contact required for the interventions, the blinding of participants and intervention staff was not feasible. The impossibility of blinding intervention staff may have led to performance bias. Only three studies (Ayele et al., 1993; Ginsburg et al., 2014; Omer et al., 2008) reported the blinding of outcome assessors. While Chang et al. (2011) measured biological outcomes and Andreoni et al. (2016) used GPS-enabled phones meaning blinding of outcome assessors may not have been entirely crucial, the more subjective outcome collection tools used in the remaining studies (i.e. observational service tests, client questionnaires, household interviews, direct observations of CHW performance, and data collected by CHWs) make detection bias a notable risk.

Incomplete outcome data (attrition bias)

Final attrition rates for the group about which primary outcome data were gathered ranged from 0% (Ayele et al., 1993) to 30% (Derenzi et al., 2012C). Carrón et al. (1994) reported that an unspecified level of provider mortality unevenly affected the 180 CHWs randomised at baseline and attempted to compensate for this by randomly

For peer review only

deleting data from arms of the trial with less attrition until group numbers were even. Not only does this process obscure the actual attrition rate, it still leaves the study at high risk of attrition bias: randomly deleting data is not the same as never having collected it; CHWs missing from one group as a result of death may be systematically different than the ones who were randomly deleted from the other groups. Three of the studies that collected data from patients did not identify those patients a priori or include data on CHW dropout so rates of attrition are unclear (Gautham et al., 2015; Ginsburg et al., 2014; Winch et al., 2003). One of the studies, however, specified that analysis nonetheless followed the intention to treat principle, in the sense that all women in the intervention communities were analysed as though exposed, whether or not CHWs had used the intervention materials while treating them (Omer et al., 2008).

Selective reporting (reporting bias)

The variety of measurements employed in the fourteen trials suggests that no standard set of outcomes exists for the evaluation of CHW performance. Data were reported in several of formats; many were unsuitable for meta-analysis even had it been advisable. Repeated attempts were made to contact trialists for missing statistical data; while many authors were helpful in providing additional data, further information was often unavailable due to data loss or non-response. Of the trials, only two reported having a protocol (Andreoni et al., 2016; Chang et al., 2011). While Chang et al. (2011) reported all pre-specified outcomes, the outcomes listed in the protocol for the Andreoni et al. (2016) study were embargoed. As with Andreoni et al. (2016) the risk of selective reporting among the other eleven trials is thus unclear.

Publication bias

For peer review only

Despite an exhaustive literature search, it is possible that the findings are subject to publication bias. Funnel plots were not used to explore publication bias as there were too few studies to draw conclusions from such analysis.

C-RCTs: Recruitment bias

No C-RCT recruited individuals to clusters post-randomisation, so risk of recruitment bias across trials is low.

C-RCTs: Baseline imbalance

Three studies (Ashraf et al., 2015; Carrón et al., 1994; Omer et al., 2008) did not provide an explicit statement of group equivalence at baseline, though Carrón et al. (1994) pair-matched clusters to reduce the risk of such differences. Two studies reported baseline differences: Ayele et al., (1993) reported baseline differences in CHW age and years since graduation but did not control for them; despite pair-matching clusters. Winch et al. (2003) reported baseline differences in literacy, spouse's education, and carer's education and adjusted for economic status (as measured by presence of a metal roof), education, and literacy in a multivariate random effect logistic regression model. Chang et al. (2011) and Bossuroy et al. (2016) reported that there were no significant baseline differences among participants.

C-RCT: Loss of clusters

Ashraf et al. (2015) and Bossuroy et al. (2016) did not note whether clusters were lost; in the remaining five studies, no clusters were lost.

C-RCT: Unit of analysis issues

Only three studies (Ashraf et al., 2015; Bossuroy et al., 2016; Chang et al., 2011; Winch et al., 2003) explicitly took clustering into account in the analysis, the rest were unclear or did not.

For peer review only

C-RCTs: Comparability with nonrandomised trials

Because no meta-analysis was conducted, comparability with individually randomised trials wasn't relevant.

C-RCT Contamination

Contamination and threats to programme differentiation were unlikely given the geographical separation of the different clusters and the fact that the experimental and control interventions were delivered by different personnel.

Appendix H: Cost Data

Carrón et al. (1994), Chang et al. (2011), and Derenzi et al. (2012) reported cost outcomes. Though cost was not prespecified in the review protocol as a variable about which information would be collected, it is important in assessing the relative merits of various combinations of inputs to improve CHW performance and so it is reported here Because the data on resource use and costs were very heterogeneous, meta-analysis was not appropriate and available data are presented in the table below

The economics component of this review was conducted according to current guidance on the use of economics methods in the preparation and maintenance of Cochrane reviews (Higgins & Green, 2011). The reviewer classified the included economic evaluations based on an established system (Drummond et al., 2005). Briefly, full economic evaluations (e.g. cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis) compare the costs (resource use) associated with one or more alternative interventions (e.g. intervention X versus comparator Y) with their consequences (outcomes, effects) whereas partial economic evaluations (e.g. cost analyses, cost-description studies, and cost-outcome descriptions) do not make explicit comparisons

For peer review only

between alternative interventions in terms of both costs (resource use) and consequences (effects) (Higgins & Green, 2011).

Table Summary of cost data

Primary trial (Cost paper)	Type of economic evaluation (Perspective, horizon)	Study pop.	Intervention vs. Comparator	Economic results (Currency & price year)
Carrón et al. 1994	CA (Provider, 1 year)	CHWs	I_a = Decision tree I^b = Behavioural feedback instrument $I_{a+b} = I_a + I^b$ C = Usual care	I_a was the least expensive (29.42 USD/CHW) followed by I^b (78.45 USD/CHW), the I_{a+b} (106.80 USD/CHW), and C, the traditional training course (166.17 USD/CHW in groups of twenty-five, 118.40 USD/CHW in groups of fifteen or fewer) (1993 USD)
Chang et al. 2011	CA (Provider, 2 years)	HIV positive adults	I = A text message reporting system and health centre warmline C = Usual care	The total cost of the mHealth intervention was $2,353.24. The yearly cost was $1045.88, resulting in a yearly additional cost per patient of $2.35. (2006-08 US)
DeRenzi et al., 2012B	CA (Provider, 1 year)	Chronic care patients	I = Escalating text message reminder system & follow-up by supervisor C = Usual care	The intervention adds an estimated $0.84 per client per year of running the ICT4CHW programme with automated escalating reminders (Year unspecified USD)

CA: Cost analysis; **USD**: US dollars.

These studies generally highlight the low incremental costs or potential cost savings of these CHW performance interventions. The quality of the economic evaluations was assessed using the Cochrane-recommended CHEC list (Evers et al., 2005), a consensus criteria list that incorporates items from Drummond and Jefferson (1996) and other systematically-identified guidelines. These assessments are reported in the table below. Because CHEC was designed to assess full economic evaluations, items not applicable to the partial economic evaluations included are shaded grey. There was some risk of bias across the studies: not all relevant alternatives and costs (e.g. training, supervision) were reported in all studies, studies did not perform discounting, and no studies adequately addressed ethical and distributional implications.

Table Quality of included economic evaluations

	Item	Carrón 1994	Chang 2011	DeRenzi 2012ABC
1	Is the study population clearly described?	+	+	+
2	Are competing alternatives clearly described?	+	+	+
3	Is a well-defined research question posed in answerable form?	-	-	-
4	Is the economic study design appropriate to the stated objective?			
5	Is the chosen time horizon appropriate to include relevant costs and consequences?			
6	Is the actual perspective chosen appropriate?	+	+	+
7	Are all important and relevant costs for each alternative identified?	+	+	-

For peer review only

8	Are all costs measured appropriately in physical units?	+	+	+
9	Are costs valued appropriately?	+	+	+
10	Are all important and relevant outcomes for each alternative identified?			
11	Are all outcomes measured appropriately?			
12	Are outcomes valued appropriately?			
13	Is an incremental analysis of costs and outcomes of alternatives performed?			
14	Are all future costs and outcomes discounted appropriately?	-	-	-
15	Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	-	-	-
16	Do the conclusions follow from the data reported?	+	+	+
17	Does the study discuss the generalizability of the results to other settings and patient/ client groups?			
18	Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?	+	+	-
19	Are ethical and distributional issues discussed appropriately?	-	-	-

References for Appendix H

Higgins, J. P. T., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*: The Cochrane Collaboration.

Drummond, M., & Jefferson, T. (1996). Guidelines for authors and peer reviewers of economic submissions to the BMJ. *British Medical Journal, 313*(7052), 275-283.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford, UK: Oxford University Press.

Evers, S., Goossens, M., de Vet, H., van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *Int J Technol Assess Health Care, 21*(2), 240-245.

For peer review only

Appendix I: PRISMA Checklist

SECTION/TOPIC	#	CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	2
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Appendix A
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4-5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-6 and Appendix C
Search	8	Present full electronic search strategy for at least one database, including any	5-6 and Appendix C

For peer review only

		limits used, such that it could be repeated.	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6-7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6-7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Appendix D
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6-7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	6-7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7-8
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period)	9-13 and Appendix E

For peer review only

		and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Appendix E & G
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	14-22
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	14-22
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	22
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	23-27
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	23-27
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	27

For peer review only

PRISMA Checklist

SECTION/TOPIC	#	CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	2
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Appendix A
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4-5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-6 and Appendix C

For peer review only

Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5-6 and Appendix C
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6-7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6-7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Appendix D
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	6-7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	N/A
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	6-7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7-8

For peer review only

Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	9-13 and Appendix E
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Appendix E & G
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	14-22
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	14-22
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	22
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	23-27
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	23-27
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	27

For peer review only