BMJ Open

Systematic review of interventions for improving the performance of community health workers in low-income and middle-income countries

Madeleine Ballard,[1] Paul Montgomery[2]

[1]Centre for Evidence-Based Intervention, University of Oxford, Oxford, UK
[2]Department of Social Policy and Social Work, University of Birmingham, Birmingham, UK

Correspondence to
Madeleine Ballard;
madeleine.ballard@gmail.com

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 Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.

Methods 19 electronic databases were searched with a highly sensitive prespecified strategy and the grey literature examined, completed July 2016. Randomised controlled trials evaluating interventions to improve CHW performance in low- and middle-income countries were included and appraised for risk of bias. Outcomes were biological and behavioural patient outcomes (primary), use of health services, quality of care provided by CHWs and CHW retention (secondary).

Results Two reviewers screened 8082 records; 14 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 (1) improving behavioural outcomes for patients, (2) improving use of services by increasing the absolute number of patients who access services and, perhaps, better identifying those who would benefit from such services and (3) 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.

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 Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.

► The review strengths include its prespecified, 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 non-response 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 programme of the past, the review is timely.

INTRODUCTION

Community health workers (CHWs)—lay workers to whom simple medical procedures can be 'task shifted' from higher level medical providers (eg, 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 programme often failed to replicate the success of smaller community-based programmes.[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 programmes, however, remain unfavourable.[20–22] Yet, it is difficult to assess on the basis of such evaluations whether a given CHW programme did not achieve a statistically significant effect because (1) such programme do not work, (2) the CHW programme was not implemented properly (type III error) or (3) the CHW programme design was not yet optimised to achieve maximum effect (eg, 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 suboptimal programme design,[2][27–30] a systematic review of studies disaggregating the effect of individual intervention components could help improve programme 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 programme components (eg, recruitment, supervision)—can be used to optimise CHW performance.

METHODS
Inclusion criteria
Randomisation allows the effects of particular programme components to be isolated; this is a critical step in intervention optimisation and the goal of this review (see protocol in online supplementary appendix A).[24] Because health workforce outcomes are empirically testable and workplace interventions possible to randomise, eligible trials were randomised controlled trials (RCTs), cluster-randomised trials, cross-over 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 (eg, providing antiretroviral therapy (ART) adherence support) in the context of different programme designs (eg, more or less supervision) that have to date been excluded from reviews of CHW effectiveness (eg, 10).

Several definitions for and variants of the term CHW have been employed in the literature. For the purpose of this review, CHWs were defined as: any lay health workers who (1) live in the catchment they serve, (2) are primarily based in the community (as opposed to a health facility), (3) belong to the formal health system (ie, are managed by the government or an implementing non-governmental organisation), (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 online supplementary appendix B), there were no restrictions on the types of patients for whom data were extracted.

Any intervention designed to improve CHW performance, compared with CHWs who did not experience the intervention, was included. Studies that were designed to examine the efficacy of a particular therapy (eg, zinc tablets for diarrhoea), as opposed to the relative effectiveness of different ways of designing programme 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), use 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:

1. 3ie Impact Evaluation Database
2. Allied and Complementary Medicine Database
3. BiblioMap
4. British Library for Development Studies at IDS
5. Cumulative Index to Nursing and Allied Health Literature
6. Central Register of Controlled Trials
7. Database of Promoting Health Effectiveness Reviews
8. Embase
9. Global Health
10. IDEAS/RePEc: Economics and Finance Research Database
11. Innovations for Poverty Action Database
12. Databases within Institute for Scientific Information Web of Science
13. JOLIS World Bank-International Monetary Fund Library
14. Abdul Latif Jameel Poverty Action Lab Evaluations Database
15. Medline
16. Medline in-process and other non-indexed citations
17. Population Information Online
18. PubMed
19. World Bank Independent Evaluation Group
20. WHO Library Database

The following strategy was refined in consultation with an information specialist and used to search Medline. 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 online supplementary appendix C).

1. ((community or village) adj2 (agent$ or aid$ or promot$ or mobili?er$ ordistribut?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$ ORbrigadista$ or kader$ or monitora$ or promotora$ or sevika$ or fhw$ orchw$ or lhw$ or vhw$ OR chv$ or "shastho shebika" or "shasto karmis" oranganwadi or "barefoot doctor" OR "agente comunitario de salud" or "agentecommunitario 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 4 May 2014 and updated ending 12 July 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 Effective Practice and Organisation of Care Group data collection checklist,[34] the Template for Intervention Description and Replication[35] and the Oxford Implementation Index[36] (see online supplementary 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 (ie, multiple studies with similar interventions and outcomes) was assessed using the Grading of Recommendations, Assessment, Development and Evaluation (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 (ie, 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 (figure 1). A total of 12 records reporting 14 studies were included.[42–53] Complete information for each trial is provided (online supplementary

appendix E); study characteristics are described in table 1.[35] Information on excluded, ongoing and studies awaiting classification is provided in online supplementary appendix F.

Ten countries were represented: Bangladesh, Ethiopia, Guatemala, India (2), Mali, Pakistan (2), Paraguay, Uganda, Tanzania (3), and 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 programme component.[47–51] No information about implementation fidelity—that is, 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 online supplementary appendix A and are considered when assessing the findings.

Effects of interventions

Results are presented in a structured summary.[36] Where possible, standardised measures of effect (risk ratio (RR), standardised mean (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 online supplementary 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 are considered in detail below.

Only one study reported biological outcomes for patients,[48] finding no significant differences in HIV patients' cumulative risk of virological failure (RR=1.17, 95% CI 0.84 to 1.64, p=0.34) or mortality (RR=0.82, 95% CI 0.55 to 1.22, p=0.33).
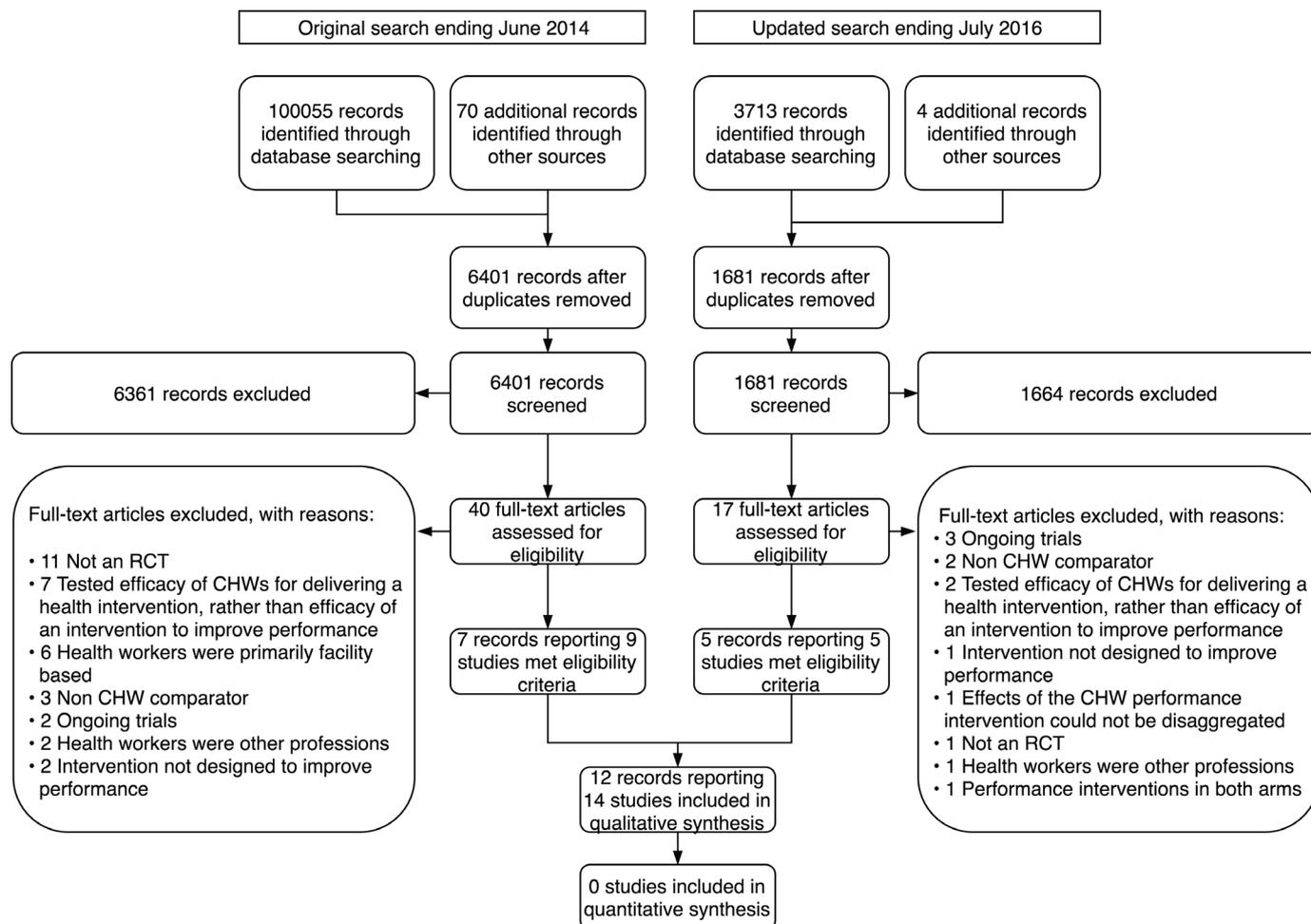
Figure 1 PRISMA study flow diagram. CHW, community health workers; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT, randomised controlled trials.

Six studies reported behavioural outcomes for patients: Bossuroy[53] found no significant difference in patient tuberculosis (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 to 1.06, p=0.59) or loss to follow-up (RR=1.29, 95% CI 0.50 to 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 to 2.41, p<0.0001), give colostrum to their newborn babies (RR=1.21, 95% CI 1.12 to 1.32, p<0.0001) and maintain exclusive breastfeeding for 4 months but equally likely to stop routine heavy work.[i] No quantitative data were 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 to 1441.23, p<0.0001). DeRenzi (43B) 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 (43C), control CHWs who had been in the short messaging service (SMS)+Supervision group during DeRenzi (43B) 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 (43B) 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).

Secondary outcomes

Secondary outcomes were use of health services, quality of care provided by CHWs and CHW retention (table 3). Overall, these studies suggest that CHW performance interventions can improve use 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

---

[i] 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 to 1.422, p<0.0001).

Table 1 Brief study characteristics

Study	Setting (study design)	CHW and patient sample	Intervention type and 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 and urban Pakistan (2×2 factorial RCT)	Female CHWs (I_{a+c}=85, I_{a+d}=84, C_{b+c}=88, C_{b+d}=80). Patients: children under 5 years. (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 years and women (n=unspecified)	**Recruitment** I: Recruitment materials emphasising career possibilities C: Recruitment materials emphasising benefits to community	**CHWs:** Home visits, patients seen, community meetings organised **Patients:** Use of services (18 months)
Ayele 1993[51] Objective: test the effect of an in-person refresher course and monthly 'supportive' supervision on CHW-led 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[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 (four-arm C-RCT)	CHWs (n=78). Adult patients (n=2500).	**Incentives** I_a: Financial incentives based on patient detection for 6 months then based on treatment adherence I_b: Financial incentives based on patient detection for 6 months then a fixed salary I_c: Fixed salary for 6 months then incentives based on treatment adherence C: Fixed salary	**CHWs:** TB case detection **Patients:** TB default rate (6 months)

Continued

Table 1 Continued

Study	Setting (study design)	CHW and patient sample	Intervention type and description of trial arms	Outcomes (follow-up)
Carrón 1994[50] Objective: test the effect of an interactive counselling guide and supervisory behavioural feedback instrument on CHWs distributing family planning materials	Rural Paraguay (2×2 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[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:** virological 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 and supervisor follow-up for underperforming CHWs C: No reminders	**CHWs:** none **Patients:** mean time overdue at clinic (40 days)
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 and 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)

Table 1 Continued

Study	Setting (study design)	CHW and patient sample	Intervention type and description of trial arms	Outcomes (follow-up)
Ginsburg 2014[44] Objective: test the effect of mobile guidance applications on CHWs referring women with abnormal breast exams	Rural Bangladesh (three-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 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[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 years. (n=152, C=134)	**Integration-Training-Reporting** I: Visual aids, refresher training, a counter referral record-keeping notebook, a community meeting C: Standard training	**CHWs:** proportion of children treated at home **Patients:** closed referrals

$\bar{x}$, mean; C, control; CHW, community health worker; C-RCT, cluster RCT; I, intervention; N, intervention/control participant breakdown unspecified, total number of participants; RCT: randomised controlled trial; SMS, short messaging service (ie, text message); TB, tuberculosis; yrs, years.

Table 2 Primary outcomes

Study (N)	Intervention type	Variables	Measure of effect (95% CI or SE)
Biological outcomes for patients			
Chang 2011[48] (n=970 patients)	Reporting-Supervision	Virological failure	RR=1.17 (0.84 to 1.64) p=0.34
		Mortality	RR=0.82 (0.55 to 1.22) p=0.33
Behavioural outcomes for patients			
Bossuroy 2016[53] (n=2500 patients)	Incentives	TB default rate	B=−0.04 (SE=0.06) p=0.25
Chang 2011[48] (n=970 patients)	Reporting-Supervision	<95% pill adherence	RR=1.01 (0.97 to 1.06) p=0.59
		Loss to follow-up	RR=1.29 (0.50 to 13.32) p=0.60
DeRenzi 2012B[43] (n=unspecified patients)	Supervision	Mean days clients overdue	U=271.00 p<0.001 r=0.500
DeRenzi 2012C[43] (n=unspecified patients)	Supervision	Mean days clients overdue	See text
Omer 2008[47] (n=1070 patients)	Integration-Training	Attend at least one prenatal visit	RR=1.94 (1.56 to 2.41) p<0.0001
		Colostrum to newborns	RR=1.21 (1.12 to 1.32) p<0.0001
		Stop routine heavy work	See text
		Exclusive breastfeeding	See text
Winch 2003[49] (n=286 patients)	Training-Reporting-Integration	Proportion of malarial children treated by mothers exactly per IMCI norms	RR=27.76 (0.53 to 1441.23) p<0.0001

B, regression coefficient; C, control; I, intervention; IMCI, Integrated Management of Childhood Illness; r, regression coefficient; RR, risk ratio; TB, tuberculosis; U, Mann-Whitney U-statistic.

care provided. The body of evidence for both outcomes was graded moderate (downgraded for design and implementation as allocation concealment, random sequence generation and blinding of assessors was 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 online supplementary appendix H. The data for use and quality of care is considered in detail below.

Four studies reported on use; three examined the effect of CHW performance interventions on closed referrals (43A, 44, 49) and one on use 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 three times greater in the intervention arms (I_A=3.1%, I_B=3.2%) than the control arm (1.0%) (p<0.0001). DeRenzi (43A), 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 to 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 children under five undergoing health checks by 24% (n=1618 patients, B=318.1, SE=98.05, p<0.001), the number of children under five being weighed by 23% (n=1610 patients,

Table 3 Secondary outcomes

Study	Intervention type	Variables	Measure of effect (95% CI or SE)
Use of health services			
Ashraf 2015[52] (n=specified in text)	Recruitment	Use of health services	See text
DeRenzi 2012A[43] (n=11 CHWs)	Supervision	Closed referrals per CHW	SMD=0.86 (−0.30 to 2.03)
Ginsburg 2014[44] (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[49] (n=79 patients)	Training-Reporting-Integration	Closed referrals	RR=0.78 (0.57 to 1.07) p=0.13
CHW quality of care			
Andreoni 2016[46] (n=337 CHWs)	Incentives	Closeness to vaccine distribution policy objective	SMD=−0.32 (−0.54 to −0.1)
Ashraf 2015[52] (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
		Community meetings organised	B=17.06 (SE=5.22) p<0.001
Ayele 1993[51] (n=100 CHWs)	Training-Supervision	Completion of government-set CHW activities	See text
Bailey 1996[45] (n=49)	Training	Knowledge score	Combined I: 36.8% C: 24.9% p>0.05
Bossuroy 2016[53] (n=78 CHWs)	Incentives	TB case detection rate	See text
Carrón 1994[50] (n=180 CHWs)	Training-Supervision	Quality of care score	See text
Gautham 2015[42] (n=16 CHW)	Equipment	Protocol compliance	See text
Winch 2003[49] (n=102 CHWs)	Training-Reporting-Integration	Appropriate drug administration Dispensing correct amount of medication	RR=1.45 (0.99 to 2.13) p=0.04 RR=1.22 (0.69 to 2.18) p=0.47

B, regression coefficient; C, control; CHWs, community health workers; I, intervention; RR, risk ratio; SMD, standardised mean difference; TB, tuberculosis.

B=284.3, SE=110.2, p=0.005) and number of children under one receiving immunisation 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 bacille Calmette-Guerin (BCG) vaccinations and children under one 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 (eg, 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=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), an RR 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 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 14 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 (1) improving certain behavioural outcomes for patients, (2) improving use 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 (3) 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 (eg, technical competence). The evidence for biological outcomes was assessed as moderate quality, though the lack of effect on virological 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, use 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] and (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 14 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 judgements. 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 should provide information about

Ballard M, Montgomery P. *BMJ Open* 2017;**7**:e014216. doi:10.1136/bmjopen-2016-014216

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 undersupported 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 (43ABC). 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 programme components (eg, 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 programme-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 programme 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 (43ABC) 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 programme 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] for example, in Bossuroy[53] 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[51] study faced cultural barriers working with women in the home, and so on, 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, for example, one study indicated that equipment may not be as valuable in the absence of adequate training.[44]

Limitations

First, 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. Second, 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. Third, poor reporting and non-response 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.

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 (see the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Checklist in online supplementary appendix I).

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.

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 PA, Gwatkin DR, Burger SE. Community-based health workers: head start or false start towards health for all? *Soc Sci Med* 1987;25:443–59.
3. Frankel S. *The community health worker: effective programmes for developing countries*: Oxford University Press, 1992.
4. Heggenhougen K. *Community health workers: the Tanzanian experience*: Oxford University Press., 1987.
5. Walt G. CHWs: are national programmes in crisis? *Health Policy Plan* 1988;3:1–21.
6. Gilson L, Walt G, Heggenhougen K, *et al*. National community health worker programs: how can they be strengthened? *J Public Health Policy* 1989;10:518.
7. Walt G. *Community health workers in national programmes: just another pair of hands? 1990, Milton Keynes*. UK: Open University Press.
8. Kredo T, Adeniyi FB, Bateganya M, *et al*. Task shifting from doctors to non-doctors for initiation and maintenance of antiretroviral therapy. *Cochrane Database Syst Rev* 2014;7:CD007331.
9. Lassi ZS, Haider BA, Bhutta ZA. Community-based intervention packages for reducing maternal and neonatal morbidity and mortality and improving neonatal outcomes. *Cochrane Database Syst Rev* 2010:CD007754.
10. Lewin S, Munabi-Babigumira S, Glenton C, *et al*. Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases. *Cochrane Database Syst Rev* 2010:CD004015.
11. Okwundu CI. Home- or community-based programmes for treating malaria. *Cochrane Database of Syst Rev* 2013;5:CD009527.
12. United Nations. The Millennium Development Goals Report 2015. 2016 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. *Bull World Health Organ* 2005;83:5–6.
14. Kruk ME, Myers M, Varpilah ST, *et al*. What is a resilient health system? Lessons from Ebola. *Lancet* 2015;385:1910–2.
15. World Health Organization, 2008. *The Addis Ababa Declaration on task shifting. in The first international conference on task shifting*, WHO
16. Campbell J. *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. *BMJ* 2011;342:d696.
18. Ghani R. 40 Million children have no access to the most basic healthcare services. *BMJ* 2011;343:d4667.
19. Singh P, Sachs JD. 1 million community health workers in sub-Saharan Africa by 2015. *Lancet* 2013;382:363–5.
20. Munos M, Guiella G, Roberton T, *et al*. Independent Evaluation of the Rapid Scale-Up Program to Reduce Under-Five Mortality in Burkina Faso. *Am J Trop Med Hyg* 2016;94:584–95.
21. Amouzou A, Hazel E, Shaw B, *et al*. Effects of the integrated Community Case Management of Childhood Illness Strategy on Child Mortality in Ethiopia: A Cluster Randomized Trial. *Am J Trop Med Hyg* 2016;94:596–604.
22. Amouzou A, Kanyuka M, Hazel E, *et al*. Independent Evaluation of the integrated Community Case Management of Childhood Illness Strategy in Malawi Using a National Evaluation Platform Design. *Am J Trop Med Hyg* 2016;94:574–83.
23. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent* 2011;71:52–63.
24. Collins LM, Baker TB, Mermelstein RJ, *et al*. The multiphase optimization strategy for engineering effective tobacco use interventions. *Ann Behav Med* 2011;41:208–26.
25. Hasson H. Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implement Sci* 2010;5:11.
26. Basch CE, Sliepcevich EM, Gold RS, *et al*. Avoiding type III errors in health education program evaluations: a case study. *Health Educ Q* 1985;12:315–31.
27. Alam K, Khan JA, Walker DG. Impact of dropout of female volunteer community health workers: an exploration in Dhaka urban slums. *BMC Health Serv Res* 2012;12:260.
28. Olang'o CO, Nyamongo IK, Aagaard-Hansen J. 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):232–7.
29. Medhanyie A, Spigt M, Dinant G, *et al*. Knowledge and performance of the Ethiopian health extension workers on antenatal and delivery care: a cross-sectional study. *Hum Resour Health* 2012;10:44.
30. Puchalski Ritchie LM, van Lettow M, Barnsley J, *et al*. Evaluation of lay health workers' needs to effectively support anti-tuberculosis treatment adherence in Malawi. *Int J Tuberc Lung Dis* 2012;16:1492–7.
31. Standing H, Chowdhury AM. Producing effective knowledge agents in a pluralistic environment: what future for community health workers? *Soc Sci Med* 2008;66:2096–107.
32. World Health Organization, 1987. *Community health workers: pillars for health for all report of the Interregional conference*, Yaoundé, CameroonUnpublished document SHS/CIH/87.2
33. Khan KS. *Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews*: NHS Centre for Reviews and Dissemination, 2001.
34. McAuley L, Ramsay C. Cochrane Effective Practice and Organisation of Care Review Group (EPOC) Data Collection Checklist. 2014 htt

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

35. Hoffmann TC, Glasziou PP, Boutron I, *et al*. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348:1687.

36. Montgomery P, Underhill K, Gardner F, *et al*. The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *J Clin Epidemiol* 2013;66:874–82.

37. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. 2011*: The Cochrane Collaboration.

38. Guyatt GH. Rating quality of evidence and strength of recommendations: GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924.

39. Rehfuess EA, Akl EA. Current experience with applying the GRADE approach to public health interventions: an empirical study. *BMC Public Health* 2013;13:9.

40. Waters E, Hall BJ, Armstrong R, *et al*. Essential components of public health evidence reviews: capturing intervention complexity, implementation, economics and equity. *J Public Health* 2011;33:462–5.

41. Movsisyan A, Melendez-Torres GJ, Montgomery P. Users identified challenges in applying GRADE to complex interventions and suggested an extension to GRADE. *J Clin Epidemiol* 2016;70.

42. Gautham M, Iyengar MS, Johnson CW. Mobile phone-based clinical guidance for rural health providers in India. *Health Informatics J* 2015;21:253–66.

43. Derenzi B, 2012. *Improving Community Health Worker Performance Through Automated SMS ACM International Conference Proceeding Series*

44. Ginsburg OM, Chowdhury M, Wu W, *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;19:177–85.

45. Bailey JE, Coombs DW. Effectiveness of an Indonesian model for rapid training of Guatemalan health workers in diarrhea case management. *J 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, Mhatre S, Ansari N, *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.

48. Chang LW, Kagaayi J, Arem H, *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 Behav* 2011;15:1776–84.

49. Winch PJ, Bagayoko A, Diawara A, *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. *Trans R Soc Trop Med Hyg* 2003;97:481–90.

50. Carrón J, Melian M, Leon F. 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, Desta A, Larson C. The functional status of community health agents: atrial of refresher courses and regular supervision. *Health Policy Plan* 1993;8:379–84.

52. Ashraf N, Bandiera O, Lee SS. Do-Gooders and Go-Getters: career incentives, selection, and performance in public service delivery. 2015.

53. Bossuroy T, Delavallade C, Pons V. 2016. Fighting tuberculosis through community based counsellors: a randomized evaluation of performance based incentives in India in 3ie Grantee Final Report. New Delhi: International Initiative for Impact Evaluation.

54. Kok MC, *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.

55. Naimoli JF, Frymus DE, Wuliji T, *et al*. A Community Health Worker "logic model": towards a theory of enhanced performance in low- and middle-income countries. *Hum Resour Health* 2014;12:56.

56. Schünemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 13. Applicability, transferability and adaptation. *Health Res Policy Syst* 2006;4:25.

57. Oxman AD, Lavis JN, Fretheim A, *et al*. SUPPORT Tools for evidence-informed health Policymaking (STP) 16: Using research evidence in balancing the pros and cons of policies. *Health Res Policy Syst* 2009;7():16.

58. Benabou R, Tirole J. Intrinsic and Extrinsic Motivation. *Rev Econ Stud* 2003;70:489–520.

59. Kok MC, Dieleman M, Taegtmeyer M, *et al*. Which intervention design factors influence performance of community health workers in low- and middle-income countries? A systematic review. *Health Policy Plan* 2015;30:1207–27.