WHO guideline on health policy and system support to optimize community health worker programmes

Annex 6. Summary of evidence

This web annex contains a tabulated summary of the evidence gathered to inform development of the *WHO guideline on health policy and system support to optimize community health worker programmes*. Fifteen systematic reviews were undertaken to assess the evidence on the policy questions specifically examined in the guideline. Each review considered a specific question related to the effectiveness of community health worker (CHW) programmes, and was structured according to the standard population, intervention, control, outcome (PICO) approach. For each PICO question, the summary of evidence presented below includes, as applicable:

a. overview of included quantitative studies

b. GRADE quality assessment

c. Newcastle-Ottawa quality assessment for cohort studies and cross-sectional studies

d. qualitative findings

e. evidence to decision tables.

Note: There are some variations in the way the evidence is presented in the tables for each section, due to differences in the availability, quantity and quality of the evidence, and in the approach adopted by different authors of the systematic reviews.

Contents

Abbreviations, Annex 6

ART	antiretroviral treatment/therapy
ASHA	accredited social health activist
CC	correlation coefficient
CHA	community health adviser/agent/assistant
CHV	community health volunteer
CHW	community health worker
CI	confidence interval
CTA	community treatment assistant
DALY	disability-adjusted life-year
DOTS	directly observed treatment, short course
DTP	diphtheria-tetanus-pertussis
EM	enhanced management
EPT	efficient product transport
FHW	female health worker
GDG	Guideline Development Group
HEW	health extension worker
iCCM	integrated community case management
LHW	lady health worker
MD	mean difference
MoTeCH	Mobile Technology for Community Health
OR	odds ratio
PHW	peer health worker
PICO	population, intervention, control, outcome
RCT	randomized controlled trial
RDT	rapid diagnostic test
RR	relative risk / risk ratio
SMD	standardized mean difference
TB	tuberculosis
TT	trachomatous trichiasis
VHV	village health volunteer
VHW	village health worker
WHO	World Health Organization

6.1 *Selection.* For CHWs being selected for pre-service training, what strategies for selection of applications for CHWs should be adopted over what other strategies?

Quantitative findings: overview of included quantitative studies

Study	Design	Setting and participants	Intervention	Comparison or control	Measures, data collection	Findings	Outcomes
Observational studies							
Posner et al. *(1)*	Longitudinal cohort study	In Nepal, a combination of 504 in-school and out-of-school adolescent girls were selected and trained to be peer educators	Peer education programme	Pre-post comparison of attitudes and behaviours of the peer educators, relative to training. Comparisons between castes and regions	Structured questionnaire before and after training Main outcome variables: • knowledge on HIV and menstrual prohibitions • leadership self-efficacy • collective efficacy	Leadership self-efficacy: increased from baseline to endline ($P < 0.001$) Education influenced leadership self-efficacy: girls with only primary education scored lower than those with higher levels of education Collective efficacy: perceptions of collective efficacy increased from baseline to endline ($P < 0.001$) Knowledge: knowledge of HIV and sexually transmitted infections increased 15% from baseline to endline ($P < 0.01$) Menstrual restrictions: average number of restrictions peer educators observed decreased ($P < 0.001$)	Outputs: knowledge, competency
Kansal, Kumar and Kumar *(2)*	Cross-sectional study (survey)	135 accredited social health activists (ASHAs) located in Uttar Pradesh, India		Random sampling technique, no comparison or control	Personal interviews, data from ASHAs, stakeholders and some beneficiaries Data included ASHA socioeconomic characteristics, knowledge and practices (assessed by educational qualification), and assessment of work performance in community	Education level: class 8 (31.3%), high school (36.6%), intermediate (22.2%), graduation and above (10.4%). Education status affected ability to fill out index register; those with education up to class 8 struggled Knowledge: higher amongst those with education up to intermediate, followed by graduate and above. Majority had good knowledge of	Outputs: knowledge, productivity, competency

Study	Design	Setting and participants	Intervention	Comparison or control	Measures, data collection	Findings	Outcomes
						antenatal and child care services	
Kawakatsu et al. *(3)*	Cross-sectional study (survey)	Based in western Kenya, survey on determinants of performance of all active CHWs (*n* = 1788) in 64 community units, their supervisors, and a random sample of mothers with young children (*n* = 2560)		Multilevel modelling: • level 1: CHWs • level 2: supervisors and community level	CHW performance based on three indicators: reporting rate, health knowledge, household coverage. Assessed by three separate surveys as per different study participants	Significant factors that influenced CHW performance: • Marital status: married more likely to give higher performance • Education level: a middle and high status were positive significant factors for performance • Size of household: larger household, performance significantly increased • Work experience: longer work experience positive factor influencing performance • Personal sanitation practice: better sanitation practice, positive factor • Number of supervisions received: number received negatively impacted performance	Outputs: knowledge, productivity

GRADE quality assessment

Outcome(s)	Study design	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect	Dose–response gradient	Plausible control for confounding
ASHAs practised developing a village health plan: eighth standard vs graduate	Observational	Outputs	Indirect	Competencies	135	1	●	No	No	No	No	No	No	No
Education level associated with CHW knowledge	Observational	Outputs	Indirect	Knowledge	639	2	●	No	No	No	No	No	No	No
Higher education status associated with higher performance	Observational	Outputs	Indirect	Productivity	1923	2	●	No	No	No	No	No	No	No
Longer work experience associated with higher performance and best-practice sanitation	Observational	Outputs	Direct	Productivity	1788	1	●	No	No	No	No	No	No	No
Married CHW associated with higher performance	Observational	Outputs	Direct	Productivity	1788	1	●	No	No	No	No	No	No	No

Legend

● Low risk of bias.

● Unclear risk of bias.

● High risk of bias.

Outcome(s)	Relative effect: 95% confidence interval (CI)	Number of participants (studies)	Quality of evidence (GRADE)	Comments
ASHA practised developing a village health plan: eighth standard vs graduate	0.79 (0.65–0.93)	135 (1)	⨁ [a,b] Very low	It is uncertain whether education level is associated with CHW competence
Education level associated with CHW knowledge	0.65 (0.43–0.87)	639 (1)	⨁ [a] Very low	It is uncertain whether education level is associated with CHW knowledge
Higher education status associated with higher performance	0.93 (0.68–1.19)	1923 (1)	⨁ [a] Very low	It is uncertain whether education level is associated with CHW productivity
Longer work experience associated with higher performance	0.34 (0.13–0.55)	1788 (1)	⨁ [a] Very low	It is uncertain whether length of work experience is associated with CHW productivity
Married CHW associated with higher performance	0.74 (0.00–1.50)	1788 (1)	⨁ [a] Very low	It is uncertain whether marriage status is associated with CHW productivity

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Notes on GRADE scores

a. Downgraded one level for risk of bias: information comes from study/ies assessed as high risk of bias for the majority of domains.

b. In the context of the significant delays we have experienced in obtaining both a response and information from schools regarding module selection.

Newcastle-Ottawa quality assessment for cohort studies and cross-sectional studies

	Selection				Comparability	Outcome			
Study	**Representativeness of exposed sample**	**Selection of exposed sample**	**Ascertainment of exposure**	**Demonstration that outcome of interest was not present at start of study**	**Comparability of cohorts on basis of design or analysis**	**Assessment of outcome**	**Was follow-up long enough for outcomes to occur**	**Adequacy of follow-up of cohorts**	**Total score**
Possible score	1	1	1	1	2	1	1	1	Max: 9
Posner et al. *(1)*	✹	✹	✹				✹		4

	Selection				Comparability	Outcome		
Study	**Representativeness of sample**	**Sample size**	**Non-respondents**	**Ascertainment of exposure**		**Assessment of outcome**	**Appropriateness of statistical test**	**Total score**
Possible score	1	1	1	2	2	1	2	Max: 10
Kansal, Kumar and Kumar *(2)*	✹	✹				✹		3
Kawakatsu et al. *(3)*	✹	✹		✹		✹	✹	5

Qualitative findings

Objective	To identify, appraise, and synthesize qualitative research evidence on the barriers and facilitators to selection of CHWs for pre-service training		
Perspective	Experiences and perspectives of CHWs on their selection for pre-service training		
Included programmes	CHW programmes		
Review finding	**Overall CERQual assessment of confidence**	**Explanation of CERQual assessment**	**Studies contributing to the review**
Satisfaction: CHWs selected by community were more satisfied	Moderate confidence	This finding was graded as moderate confidence because of minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Studies: Adongo et al. *(4)*, Turinawe et al. *(5)*, Turinawe *(6)*, Strachan et al. *(7)*, Carter-Pokras et al. *(8)*, Cherrington et al. *(9)*
Competencies: People who are comfortable in talking about sensitive issues when required (e.g. sex and condoms)	Moderate confidence	This finding was graded as moderate confidence because of minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Blumenthal, Eng and Thomas *(10)*
Motivation and self-esteem: Being selected by the community engendered pride, and feelings of recognition and popularity	Moderate confidence	This finding was graded as moderate confidence because of minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Studies: Mercader et al. *(11)*, Abbey et al. *(12)*, Mukanga et al. *(13)*, Dil et al. *(14)*, Ruebush, Weller and Klein *(15)*
Attrition: Perceived to be lower when approved by community and family members	Moderate confidence	This finding was graded as moderate confidence because of minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Abbey et al. *(12)*

Evidence to decision table

Recommendations

*1A: WHO **suggests** using the following criteria for selecting community health workers for pre-service training:*

- minimum educational level that is appropriate to the task(s) under consideration;
- membership of and acceptance by the target community;
- gender equity appropriate to the context (considering affirmative action to preferentially select women to empower them and, where culturally relevant, to ensure acceptability of services by the population or target group);
- personal attributes, capacities, values, and life and professional experiences of the candidates (e.g. cognitive abilities, integrity, motivation, interpersonal skills, demonstrated commitment to community service, and a public service ethos).

Certainty of the evidence – very low. Strength of the recommendation – conditional.

*1B: WHO **suggests** not using the following criterion for selecting community health workers for pre-service training:*

- age (except in relation to requirements of national education and labour policies).

Certainty of the evidence – very low. Strength of the recommendation – conditional.

*1C: WHO **recommends** not using the following criterion for selecting community health workers for pre-service training:*

- marital status.

Certainty of the evidence – very low. Strength of the recommendation – strong.

Population: CHWs that undergo pre-service training

Intervention: selection based on predefined criteria

Factors	Decision	Explanations/comments
Magnitude of desirable and undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• Quantitative and qualitative studies found evidence of better outcomes with more educated CHWs and with community involvement in selection of CHWs • No supportive evidence was found for the use of gender, age and marital status as selection criteria
Certainty of evidence	● Very low ○ Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as very low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes

Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ○ Probably favours the intervention ○ Favours the intervention ● Varies ○ Don't know	• On balance the evidence probably favours memberships of the community and appropriate minimum education level as selection criteria • Probably does not favour age, sex and marital status as selection criteria
Resource use and cost-effectiveness	○ Large costs ● Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• Applying selection criteria may entail costs. Applying suitable selection process will probably lead to improved cost-effectiveness, uptake and acceptability and to reducing attrition. However no cost-effectiveness evidence was found on this aspect
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The Guideline Development Group (GDG) was of the view that the application of appropriate selection criteria is likely to increase health equity
Acceptability and feasibility of intervention	○ No ○ Probably no ● Probably yes ○ Yes ○ Varies ○ Don't know	• The systematic review team identified qualitative evidence suggesting that community selection improves acceptability of CHWs by the communities • The stakeholder perception survey found high levels of feasibility and acceptability of criteria such as membership of community and essential attributes of CHWs, with lower levels of acceptability and feasibility of selection based on level of education and, especially, age

Annex 6.1 references

1. Posner J, Kayastha P, Davis D, Limoges J, O'Donnell C, Yue K. Development of leadership self-efficacy and collective efficacy: adolescent girls across castes as peer educators in Nepal. Global Public Health. 2009;4(3):284–302.

2. Kansal S, Kumar S, Kumar A. Is educational level of ASHA matters for their effective functioning? A cross-sectional study in eastern Uttar Pradesh. Indian Journal of Community Health. 2012;4(1):41–4.

3. Kawakatsu Y, Sugishita T, Tsutsui J, Oruenjo K, Wakhule S, Kibosia K et al. Individual and contextual factors associated with community health workers' performance in Nyanza province, Kenya: a multilevel analysis. Health Services Research. 2015;15:442.

4. Adongo PB, Tapsoba P, Phillips JF, Tabong PT, Stone A, Kuffour E et al. The role of community-based health planning and services strategy in involving males in the provision of family planning services: a qualitative study in southern Ghana. Reproductive Health. 2013;10:36.

5. Turinawe EB, Rwemisisi JT, Musinguzi LK, de Groot M, Muhangi D, de Vries DH et al. Selection and performance of village health teams (VHTs) in Uganda: lessons from the natural helper model of health promotion. Human Resources for Health. 2015;13:73.

6. Turinawe EB. "Those were taken away and given money": power and reward expectations' influence in the selection of village health teams in rural Uganda. Rural Remote Health. 2016;16(2):3856.

7. Strachan C, Wharton-Smith A, Sinyangwe C, Mubiru D, Ssekitooleko J, Meier J et al. Integrated community case management of malaria, pneumonia and diarrhoea across three African countries: a qualitative study exploring lessons learnt and implications for further scale up. Journal of Glob Health. 2014;4(2):020404.

8. Carter-Pokras OD, Jaschek G, Martinez IL, Brown PB, Mora SE, Newton N et al. Perspectives on Latino lay health promoter programs: Maryland, 2009. American Journal of Public Health. 2011;101(12):2281–86.

9. Cherrington A, Ayala GX, Amick H, Scarinci I, Allison J, Corbie-Smith G. Applying the community health worker model to diabetes management: using mixed methods to assess implementation and effectiveness. Journal of Health Care for the Poor and Underserved. 2008;19(4):1044–59.

10. Blumenthal C, Eng E, Thomas JC. STEP sisters, sex, and STDs: a process evaluation of the recruitment of lay health advisors. American Journal of Health Promotion. 1999;14(1):4–6.

11. Mercader HFG, Kyomuhangi T, Buchner DL, Kabakyenga J, Brenner JL. Drugs for some but not all: inequity within community health worker teams during introduction of integrated community case management. BMC Health Services Research. 2014;14(Suppl. 1):S1.

12. Abbey M, Bartholomew LK, Nonvignon J, Chinbuah MA, Pappoe M, Gyapong M et al. Factors related to retention of community health workers in a trial on community-based management of fever in children under 5 years in the Dangme West district of Ghana. International Health. 2014;6:99–105.

13. Mukanga D, Tibenderana JK, Kiguli J, Pariyo GW, Waiswa P, Bajunirwe F et al. Community acceptability of use of rapid diagnostic tests for malaria by community health workers in Uganda. Malaria Journal. 2010;9:203.

14. Dil Y, Strachan D, Cairncross S, Korkor AS, Hill Z. Motivations and challenges of community-based surveillance volunteers in the Northern region of Ghana. Journal of Community Health. 2012;37(6):1192–8.

15. Ruebush TK, Weller SC, Klein RE. Qualities of an ideal volunteer community malaria worker: a comparison of the opinions of community residents and national malaria service staff. Social Science and Medicine. 1994;39(1):123–31.

6.2 *Duration of pre-service training.* For CHWs receiving pre-service training, should the duration of training be shorter versus longer?

Quantitative findings: overview of included quantitative studies

Study	Design	Setting and participants	Intervention (trials) or training comparison (observational studies)	Comparison or control	Measures and data collection	Findings	Outcomes
Intervention studies							
Greene et al. *(1)*	Randomized controlled trial (RCT)	Villages in Kongwa, United Republic of Tanzania: community treatment assistants (CTAs) (n = 36) trained to conduct trachomatous trichiasis (TT) screening	Intervention CTAs (n = 18) received extended training of one half day duration	Control CTAs (n = 18) received standard training of 30 minutes duration	Study outcomes were the proportion of TT cases screened correctly by CTA, assessed via screening survey conducted by experienced TT grader blinded to TT status assigned by CTAs	Sensitivity of TT screening in intervention CTAs was significantly higher than control (31.2%; 95% CI, 24.9–37.6, vs 5.6%; 95% CI, 2.6–8.6, $P < 0.05$) Cases of correctly identified TT were higher in control CTAs compared to intervention; however, difference not significant (28.1% vs 15.8%, $P = 0.070$) Proportion of missed TT cases among CTAs was not significantly different between the two groups ($P = 0.269$)	Outputs: competencies
Santos et al. *(2)*	Cluster RCT	African-American community churches, Maryland, United States of America: community health advisers (CHAs) (n = 28) trained to facilitate community health education workshops	Intervention CHAs (n = 12) received training via an online platform (TB-CHA) and completed training on their own (the group took an average of 26 days to complete)	Comparison CHAs (n = 16) received training via traditional classroom teaching (TC-CHA). Training was 6 hours in duration, delivered over two sessions	Study outcomes were knowledge and comprehension of training content, assessed via the mean number of attempts required to pass course examination; and CHAs' self-perceived confidence in delivering educational workshops, assessed via post-workshop survey	Mean attempts to pass certification examination comparable between groups (TB-CHA 1.6 vs TC-CHA 1.7) Delivery of community workshops differed between groups, with TB-CHAs conducting their first workshop an average of 15 weeks post-training, compared to 7 weeks for TC-CHAs A higher proportion of CHAs in the TB-CHA group reported feeling confident to engage workshop participants (100% vs 81.3%) and respond to cancer-related questions (91.7% vs 81.3%)	Outputs: advancement

Study	Design	Setting and participants	Intervention (trials) or training comparison (observational studies)	Comparison or control	Measures and data collection	Findings	Outcomes
Harvey et al. *(3)*	Quasi-experimental	Villages in Lusaka province, Zambia: CHWs (n = 79) trained to use malaria rapid diagnostic tests (RDTs)	Intervention CHWs (n = 26) received a pictorial job aid resource plus a training session of 3 hours duration	Comparison CHWs received either the RDT manufacturer's instructions only (n = 32) (MI group) or the pictorial job aid only (n = 21) (PI group)	Study outcomes were RDT administration performance, assessed via direct observation using 16-item assessment checklist	Proportion of CHWs who correctly performed all 16 steps for conducting RDT higher in CHW groups receiving higher-intensity educational support: MI vs PI (difference 23%; 95% CI, 13–33, $P < 0.05$); PI vs PI + training (difference 10%; 95% CI, 3–17, $P < 0.05$) Accuracy of CHW RDT test interpretation (number of correct RDT readings) higher in CHW groups receiving higher-intensity educational support: MI vs PI (adjusted mean difference (MD) 26; 95% CI, 17–34, $P < 0.05$); PI vs PI + training (adjusted MD 13; 95% CI, 4–22, $P < 0.05$)	Outputs: competencies
Observational studies							
Pongvongsa et al. *(4)*	Cross-sectional	Rural villages in Lao People's Democratic Republic: village health volunteers (VHVs) (n = 137) trained to provide basic primary health care services and undertake community health surveillance activities	Number of training sessions CHWs had received: < 3 (n = 61) 3–5 (n = 45) ≥ 6 (n = 31)	NA	Study outcomes were completion by VHVs of monthly reporting of health data and service activities during the past three months, assessed via researcher-conducted interviews with VHVs, collecting data on monthly reporting, history of training, and experiences and satisfaction with the VHV job	VHVs who had received three to five training sessions (adjusted odds ratio (OR) 2.53; 95% CI, 1.08–5.93) or six or more training sessions (adjusted OR 2.84; 95% CI, 1.09–7.43) were more likely to complete their monthly reporting duties than those who received fewer than three sessions	Outputs: competencies
Furth and Crigler *(5)*	Cross-sectional	Communities in Zambia: CHWs (n = 378) were trained in providing community-based HIV/AIDS antiretroviral treatment (ART) and positive living counselling	Number of days of initial training: duration ranged from 5–14 days	NA	Study outcomes were CHW performance (task completion) in positive living and ART adherence counselling, assessed via CHW client consult assessments. Consults were audiorecorded and assessed against criteria developed using international and national guidelines for	There was no significant correlation between the number of days of initial training and CHW performance scores (correlation coefficient (CC) –0.012; P = 0.865). However, two-step cluster analysis between subpopulations (low performers vs high performers) showed that the high performers group had twice as many initial training days as the poor performers	Outputs: competencies

Study	Design	Setting and participants	Intervention (trials) or training comparison (observational studies)	Comparison or control	Measures and data collection	Findings	Outcomes
					ART adherence and positive living counselling		
Wanduru et al. *(6)*	Cross-sectional	Villages in Lira district, Uganda: CHWs (*n* = 393) trained to manage malaria, diarrhoea and pneumonia in children	Duration of CHWs initial training: 2–3 days vs 4–5 days	NA	Study outcomes were CHW knowledge and performance, assessed via a CHW-completed knowledge questionnaire based on training manuals from the Uganda Ministry of Health, and case scenarios, observed by a medical officer using standardized score checklist. Researchers set a binary score cut-off for analyses: poor performance = score < 50%; good performance = score ≥ 50%. Rationale for score cut-offs not reported	CHWs whose initial training lasted 2–3 days were more likely to have scores above 50% compared with those whose training lasted 3–5 days (CC 0.31; 95% CI, 0.12–0.80) The CHWs whose initial training lasted 2–3 days were more likely to perform better than those whose training lasted 3–5 days (adjusted OR 0.1; 95% CI, 0.04–0.41)	Outputs: competencies

Definitions

Odds ratio (OR). A measure of effect that is used to approximate relative risk (i.e., the likelihood that one group will experience the outcome given a certain exposure versus the likelihood that another group will experience the outcome given they were not exposed). When the OR is greater than 1.0, the risk is greater. When the OR is between 0 and 1, the risk is lower. When the risk is 1.0, there is no difference between groups. The further the OR is above or below 1.0, the larger the effect.

Correlation. A measure of association between two different constructs.

Significance or statistical significance. The probability that a finding was observed by chance alone. Traditionally, a finding is said to be "significant" when this probability is less than 0.05 (i.e., $P < 0.05$).

Confidence interval (CI). The estimated interval between which the measure of effect (e.g. the OR) would probably be observed if the study were conducted again on a similar sample of subjects.

Adjusted (e.g. adjusted OR). When an explanatory or causal factor's raw association with an outcome is statistically adjusted to take account of other potential explanatory factors.

GRADE quality assessment

Outcome(s)	Study design	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect (95% CI)	Dose–response gradient	Plausible control for confounding
CHW competency	RCT	Output	Indirect	Competencies	36	1	a.	🟡	Yes	No	Yes	No	No	No	No
CHW competency	Non-RCT	Output	Indirect	Competencies	79	1	b.	🔴	No	No	Yes	No	No	No	No
CHW advancement in skills	RCT	Output	Developmental	Advancement	28	1	2 years	🔴	Yes	No	Yes	No	No	No	No
CHW competency	Observational	Output	Indirect	Competencies	908	3	Range 3–13 months	🔴	Yes	No	No	No	No	No	No

Legend

🟢 Low risk of bias.

🟡 Unclear risk of bias.

🔴 High risk of bias.

Notes

a. Follow-up period not reported for studies.

b. A range was provided for the estimated magnitude of effect due to substantial heterogeneity in measures, which precluded a pooled estimate.

Outcome(s)	Impact	No. of participants (studies)	Quality of evidence (GRADE)	Comments
Competencies (RCTs)	One RCT reported three competency outcomes, of which no significant differences were found for two outcomes (cases of correctly identified TT: intervention 28.1%, vs control 15.8%, $P = 0.070$; proportion of missed TT cases: $P = 0.269$) and a positive intervention effect for one outcome (sensitivity of TT screening: intervention 31.2%; 95% CI, 24.9–37.6, vs control 5.6%; 95% CI, 2.6–8.6, $P < 0.05$)	36 (1)	⨁ [a,b,c] Very low	It is uncertain whether training of greater duration or dose improves CHW competence
Competencies (non-RCTs)	One non-RCT reported two competency outcomes. Both reported positive intervention effects. RDT: MI vs PI (difference 23%; 95% CI, 13–33, $P < 0.05$); PI vs PI + training (difference 10%; 95% CI, 3–17, $P < 0.05$). Correct RDT readings:	79 (1)	⨁ [a,c] Very low	It is uncertain whether training of greater

Outcome(s)	Impact	No. of participants (studies)	Quality of evidence (GRADE)	Comments
	MI vs PI (adjusted MD 26; 95% CI, 17–34, $P < 0.05$); PI vs PI + training (adjusted MD 13; 95% CI, 4–22, $P < 0.05$)			duration or dose improves CHW competence
Competencies (observational)	Three observational studies reported competency outcomes. One study found longer duration training associated with better competency: three to five training sessions (adjusted OR 2.53; 95% CI, 1.08–5.93) or six or more training sessions (adjusted OR 2.84; 95% CI, 1.09–7.43). One study found shorter duration associated with better competency: CHWs with 2–3 days training were more likely to have scores above 50% compared to 3–5 days training (adjusted OR 0.1; 95% CI, 0.04–0.41). One study found no association (CC –0.012; $P = 0.865$)	908 (3)	⨁ [a,e] Very low	It is uncertain whether training of greater duration or dose improves CHW competence
Advancement (RCTs)	One RCT reported three advancement outcomes, of which no significant differences were found for one outcome (mean attempts to pass certification examination: TB-CHA 1.6 vs TC-CHA 1.7), a positive intervention effect for one experimental group (TB-CHAs delivered first workshop average of 15 weeks post-training, compared to 7 weeks for TC-CHAs), and a positive intervention effect for the alternate experimental group (TB-CHA group reported feeling more confident compared to TC-CHA group to engage workshop participants – 100% vs 81.3% – and respond to cancer-related questions – 91.7% vs 81.3%)	28 (1)	⨁ [a,c,e] Very low	It is uncertain whether training of greater duration or dose improves CHW advancement

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Notes on GRADE scores

a. Downgraded one level for risk of bias: information comes from study assessed as unclear or high risk of bias for the majority of domains.

b. Downgraded one level for inconsistency: two outcomes no effect, one outcome significant effect.

c. Downgraded one level for imprecision: event rate for dichotomous outcome < 300 or sample size for continuous outcome < 400.

d. Downgraded one level for inconsistency: two studies reported positive effect, one study reported no effect.

e. Downgraded one level for inconsistency: one outcome no effect, two outcomes favoured different experimental groups.

Newcastle-Ottawa quality assessment for cross-sectional studies

Study	Selection				Comparability	Outcome		Total score
	Representativeness of sample	Sample size	Non-respondents	Ascertainment of exposure		Assessment of outcome	Appropriateness of statistical test	
Possible score	1	1	1	2	2	1	2	Max: 10
Pongvongsa et al. *(4)*	✸			✸		✸	✸	4
Smith et al. *(7)*				✸				2
Furth and Crigler *(5)*	✸			✸		✸	✸	4

Qualitative findings

Objective	To identify and synthesize qualitative evidence regarding community health worker's perceptions of the duration/dose of training received to carry out health related activities		
Perspective	Experiences and opinions of community health workers and other relevant stakeholders about community health worker training programme durations		
Included programmes	Programmes that were delivered pre-service to train any type of lay community health worker to provide community health-related services of any type to underserved populations		
Review finding	**Overall CERQual assessment of confidence**	**Explanation of CERQual assessment**	**Studies contributing to the review**
Training duration/dose linked to advancement: preference for courses longer in overall length	Moderate	This finding was graded as moderate because of minor concerns regarding coherence, moderate concerns regarding methodological limitations, and substantial concerns regarding relevance and adequacy	Wennerstrom et al. *(8)*, McLean et al. *(9)*

Evidence to decision table

Recommendation 2

*WHO **suggests** using the following criteria for determining the length of pre-service training for CHWs:*

- scope of work, and anticipated responsibilities and role;
- competencies required to ensure high-quality service delivery;
- pre-existing knowledge and skills (whether acquired through prior training or relevant experience);
- social, economic and geographical circumstances of trainees;
- institutional capacity to provide the training;
- expected conditions of practice.

Certainty of the evidence – low. Strength of the recommendation – conditional.

Population: CHWs that undergo pre-service training

Intervention: longer vs shorter duration of pre-service training

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ● Large ○ Varies ○ Don't know	• Some evidence was found suggesting that that training of greater duration or dose may be associated with improved measures of CHW competency in screening and diagnostic test performance, although with variable size of effects. Studies compared training duration of relatively short length • Qualitative data suggested that CHWs value training of greater duration • The systematic review of reviews concluded that duration of training should depend on health system context, and CHWs' pre-existing capacities and roles
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ○ Varies ● Don't know	• The systematic review did not find any studies examining any harmful or unintended consequences of variable training dose or duration
Certainty of evidence	○ Very low ● Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes

Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ● Probably favours the intervention ○ Favours the intervention ○ Varies ○ Don't know	• The GDG was of the view that a longer duration of training may be beneficial compared to a shorter duration of training, particularly in settings where CHWs have a polyvalent role, but that the appropriate duration is context specific
Resource use and cost-effectiveness	○ Large costs ● Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• The moderate costs likely to be required for longer duration of training are justified to allow CHWs to acquire the competencies they need in relation to their expected role • However, no cost-effectiveness evidence was found on this aspect
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The GDG was of the view that the longer training is likely to increase health equity through improved capacity and performance of CHWs in delivering health care to underserved communities
Acceptability and feasibility of intervention	○ No ○ Probably no ● Probably yes ○ Yes ○ Varies ○ Don't know	• The systematic review team did not identify studies assessing acceptability and feasibility of the policy option under consideration • The stakeholder perception survey found longer duration of training to be broadly acceptable and feasible

Annex 6.2 references

1. Greene GS, West SK, Mkocha H, Munoz B, Merbs SL. Assessment of a novel approach to identify trichiasis cases using community treatment assistants in Tanzania. PLoS Neglected Tropical Diseases. 2015;9(12):e0004270.
2. Santos S, Tagai E, Wang M, Scheirer M, Slade J, Holt C. Feasibility of a web-based training system for peer community health advisors in cancer early detection among African Americans. American Journal of Public Health. 2014;104:2282–9.
3. Harvey SA, Jennings L, Chinyama M, Masaninga F, Mulholland K, Bell DR. Improving community health worker use of malaria rapid diagnostic tests in Zambia: package instructions, job aid and job aid-plus-training. Malaria Journal. 2008;7:160.
4. Pongvongsa T, Nonaka D, Kobayashi J, Mizoue T, Phongmany P, Moji K. 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):1269–81.

5. Furth R, Crigler. L. Improving CHW program functionality, performance, and engagement: operations research results from Zambia. Research and evaluation report, USAID Health Care Improvement Project. Bethesda, MD: University Research Co., LLC (URC); 2012.

6. Wanduru P, Tetui M, Tuhebwe D, Ediau M, Okuga M, Nalwadda C et al. The performance of community health workers in the management of multiple childhood infectious diseases in Lira, northern Uganda: a mixed methods cross-sectional study. Global Health Action. 2016;9(1):33194.

7. Smith S, Agarwal A, Crigler L, Gallo M, Finlay A, Homsi FA. Community health volunteer program functionality and performance in Madagascar: a synthesis of qualitative and quantitative assessments. Research and evaluation report. Bethesda, MA: University Research Co. LLC (URC); 2013.

8. Wennerstrom A, Johnson L, Gibson K, Batta SE, Springgate BF. Community health workers leading the charge on workforce development: lessons from New Orleans. Journal of Community Health. 2014;39(6):1140–9.

9. McLean KE, Kaiser BN, Hagerman AK, Wagenaar BH, Therosme TP, Kohrt BA. Task sharing in rural Haiti: qualitative assessment of a brief, structured training with and without apprenticeship supervision for community health workers. Intervention. 2015;13(2):135–55.

6.3 *Competencies in curriculum for pre-service training.* For CHWs receiving pre-service training, should the curriculum address specific versus non-specific competencies?

Quantitative findings: overview of included quantitative studies

Study	Design	Setting and participants	Intervention (trials) or training comparison (observational studies)	Comparison or control	Measures and data collection	Findings	Outcomes
Intervention studies							
Bhutta et al. *(1)* 2008	Quasi-experimental (pilot trial)	Rural villages in southern Pakistan. Lady health workers (LHWs) trained to deliver community-based basic antenatal and newborn care	Intervention LHWs ($n = 96$) received training comprising the standard LHW curriculum, plus additional specific curriculum topics within community mobilization, basic newborn care and group counselling	Control LHWs (number not reported) received training comprising the standard curriculum only	Study outcomes were proportion of LHW deliveries employing selected evidence-based post-birth care practices, and stillbirth and neonatal mortality rates. Data on outcomes were collected via village household audits of births and neonatal deaths. During the audits, a structured questionnaire was used to collect information on LHW newborn care practices with women reporting a live birth in the previous 4 weeks. Data on births and newborn deaths collected via household surveys were cross-checked with data registers of health facilities	Post-intervention, the delivery of selected evidence-based post-birth care practices was proportionately higher in intervention LHWs vs controls (statistical significance not reported): • use of clean delivery kit: 63% vs 1.3% • initiation of early breastfeeding: 66.1% vs 21.1% • delaying of bathing until 6 hours after birth: 50.1% vs 30.1% Post-intervention, there was a significantly greater reduction in stillbirth and neonatal mortality rates within communities served by intervention LHWs compared to controls: • change in stillbirth rates: intervention 65.9 to 43.1 per 1000 births (risk ratio (RR) 0.66; 95% CI, 0.53–0.83, $P = 0.001$) vs control 58.1 to 60.5 per 1000 births (RR 1.04; 95% CI, 0.84–1.30, $P = 0.23$) • change in neonatal mortality rates: intervention 57.3 to 41.3 per 1000 live births (RR 0.72; 95% CI, 0.56–0.91, $P = 0.006$) vs control 52.2 to 59.8 (RR 1.14; 95% CI, 0.91–1.44, $P = 0.26$)	Outputs: service delivery Impact: mortality

Study	Design	Setting and participants	Intervention (trials) or training comparison (observational studies)	Comparison or control	Measures and data collection	Findings	Outcomes
Bhutta et al. *(2)*	Cluster RCT	Rural villages in southern Pakistan. LHWs (n = 288) trained to deliver community-based basic antenatal and newborn care	Intervention same as pilot trial: LHWs (n = 134)	Control same as pilot trial: LHWs (n = 154)	Outcome measures and data collection same as pilot trial	Post-intervention, the delivery of selected evidence-based post-birth care practices was significantly higher in intervention LHWs vs controls: • use of clean delivery kit: intervention 35%; 95% CI, 27–43 vs control 3%; 95% CI, 2–5, $P < 0.0001$ • initiation of early breastfeeding: intervention 43%; 95% CI, 33–52 vs control 27%; 95% CI, 19–36, $P = 0.03$ • delaying of bathing until 6 hours after birth: intervention 50%; 95% CI, 39–60 vs control 27%; 95% CI, 17–38, $P = 0.008$ Post-intervention, stillbirth and neonatal mortality rates were significantly lower in community clusters served by intervention LHWs compared to controls: • rates of stillbirth: intervention 39·1 per 1000 total births vs control 48·7 per 1000 total births (RR 0·79; 95% CI, 0·68–0·92, P =0·006) • rates of neonatal mortality: intervention 43·0 deaths per 1000 live births vs control 49·1 per 1000 live births (RR 0·85; 95% CI, 0·76–0·96, $P = 0·02$)	Outputs: service delivery Impact: mortality

Definitions

Risk ratio (RR). A measure of effect that is used to approximate relative risk (i.e. the likelihood that one group will experience the outcome given a certain exposure versus the likelihood that another group will experience the outcome given they were not exposed). When the RR is greater than 1.0, the risk is greater. When the RR is between 0 and 1, the risk is lower. When the RR is 1.0, there is no difference between groups. The further the RR is above or below 1.0, the larger the effect.

95% confidence interval (CI). The estimated interval between which the measure of effect (e.g. the RR) would probably be observed if the study were conducted again on a similar sample of subjects.

Significance or statistical significance. The probability that a finding was observed by chance alone. Traditionally, a finding is said to be "significant" when this probability is less than 0.05 (i.e., $P < 0.05$).

GRADE quality assessment

Outcome(s)	Study design	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect (95% CI)	Dose–response gradient	Plausible control for confounding
Mortality	RCT	Impact	CHW-attributable changes in health at the population level	Stillbirth and neonatal mortality rate	288	1	2 years	●	No	No	No	No	No	No	No
Mortality	non-RCT	Impact	CHW-attributable changes in health at the population level	Stillbirth and neonatal mortality rate	96	1	2 years	●	No	No	Yes	No	No	No	No
Service delivery	RCT	Output	Direct	Delivery of post-birth care practices in line with evidence-based recommendations	288	1	2 years	●	No	No	No	No	No	No	No
Service delivery	non-RCT	Output	Direct	Delivery of post-birth care practices in line with evidence-based recommendations	96	1	2 years	●	No	No	Yes	No	No	No	No

Legend

● Low risk of bias.
● Unclear risk of bias.
● High risk of bias.

Outcome(s)	Impact	No. of participants (studies)	Quality of evidence (GRADE)	Comments
Mortality (RCT)	One RCT reported two outcomes of CHW-attributable changes in health at the population level (mortality), with positive effects reported for both outcomes, favouring the group that received training with additional specific competencies (intervention) compared to standard training (control): stillbirth (RR 0·79; 95% CI, 0·68–0·92, *P* = 0·006) and neonatal mortality (RR 0·85; 95% CI, 0·76–0·96, *P* = 0·02)	288 (1)	⨁⨁⨁⨁ High	There is high-quality evidence from one RCT that including more specific training competencies for CHWs probably improves patient outcomes
Mortality (non-RCT)	One non-RCT reported two outcomes of CHW-attributable changes in health at the population level (mortality), with positive effects reported for both outcomes, favouring the group that received training with additional specific competencies (intervention) compared to standard training (control): stillbirth (change baseline to follow-up) (intervention RR 0.66; 95% CI, 0.53–0.83, *P* = < 0.001 vs control RR 1.04; 95% CI, 0.84–1.30, *P* = 0.23); and neonatal mortality (change baseline to follow-up) (intervention RR 0.72; 95% CI, 0.56–0.91, *P* = 0.006 vs control RR 1.14; 95% CI, 0.91–1.44, *P* = 0.26)	96 (1)	⨁ [a,b] Very low	It is uncertain whether more specific training competencies for CHWs improves patient outcomes
Service delivery (RCT)	One RCT reported three CHW service delivery outcomes (use of evidence-based birth care practices), with positive effects reported for all outcomes, favouring the group that received training with additional specific competencies (intervention) compared to standard training (control): use of clean delivery kit (intervention 35%; 95% CI, 27–43 vs control 3%; 95% CI, 2–5, *P* = < 0.0001); initiation of early breast-feeding (intervention 43%; 95% CI, 33–52 vs control 27%; 95% CI, 19–36, *P* = 0.03); and delaying of bathing till 6 hours post-birth (intervention 50%; 95% CI, 39–60 vs control 27%; 95% CI, 17–38, *P* = 0.008)	288 (1)	⨁⨁⨁⨁ High	There is high-quality evidence from one RCT that including more specific training competencies for CHWs probably improves CHW service delivery
Service delivery (non-RCT)	One RCT reported three CHW service delivery outcomes (use of evidence-based birth care practices), with positive effects reported for all outcomes, favouring the group that received training with additional specific competencies (intervention) compared to standard training (control): use of a clean delivery kit (intervention 63.0% vs control 1.3%); initiation of early breastfeeding (intervention 66.1% vs control 21.1%); and delaying of bathing till 6 hours post-birth (intervention 50.1% vs control 30.1%) (probability values not reported)	96 (1)	⨁ [a,b] Very low	It is uncertain whether more specific training competencies for CHWs improve CHW service delivery

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Notes on GRADE scores

a. Downgraded one level for risk of bias: information comes from study assessed as unclear or high risk of bias for the majority of domains.

b. Downgraded one level for imprecision: event rate for dichotomous outcome < 300.

Evidence to decision table

Recommendation 3

*WHO **suggests** including the following competency domains for the curriculum for pre-service training of CHWs, if their expected role includes such functions:*

Core

- promotive and preventive services, identification of family health and social needs and risk;
- integration within the wider health care system in relation to the range of tasks to be performed in accordance with CHW role, including referral, collaborative relation with other health workers in primary care teams, patient tracing, community disease surveillance, monitoring, and data collection, analysis and use;
- social and environmental determinants of health;
- providing psychosocial support;
- interpersonal skills related to confidentiality, communication, community engagement and mobilization;
- personal safety.

Additional

- diagnostic, treatment and care in alignment with expected role(s) and applicable regulations on scope of practice.

Certainty of the evidence – moderate. Strength of the recommendation – conditional.

Population: CHWs that undergo pre-service training

Intervention: use of specific competencies for pre-service training vs not

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• The systematic review identified two relevant studies, both conducted in Pakistan, related to the same intervention, one a pilot of a larger RCT. Both compared standardized CHW training vs such training with additional specific curriculum components on service delivery and patient (mortality) outcomes. Consistent findings were reported across the included trials, with the addition of training with specific curricula components improving CHW provision of several post-birth care practices (proportion of births) in line with evidence-based recommendations, and reducing stillbirth and neonatal mortality rates
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ○ Varies ● Don't know	• The systematic review did not find any studies examining any harmful or unintended consequences of the policy options under consideration
Certainty of evidence	○ Very low ○ Low ● Moderate ○ High ○ No included studies	• Two studies showed consistent positive effects of intervention, providing moderate evidence that the addition of more specific training competencies in the curricula improves CHW health service and patient outcomes related to that component of service provision

Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ● Probably favours the intervention ○ Favours the intervention ○ Varies ○ Don't know	• The GDG was of the view that a longer duration of training may be beneficial compared to a shorter duration of training, particularly in settings where CHWs have a polyvalent role, but that the appropriate duration is context specific
Resource use and cost-effectiveness	○ Large costs ○ Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ● Varies ○ Don't know	• The systematic review team did not identify any studies assessing resource requirements associated with the policy options of interest • No cost-effectiveness evidence was found
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The GDG was of the view that the inclusion of specific competencies in pre-service training of CHWs is likely to increase health equity through improved capacity and performance in delivering health care to underserved communities
Acceptability and feasibility of intervention	○ No ○ Probably no ● Probably yes ○ Yes ○ Varies ○ Don't know	• The systematic review team did not identify studies assessing acceptability and feasibility of the policy option under consideration • The stakeholder perception survey found high levels of acceptable and feasibility of different components of CHW training, such as preventive and promotive behaviours, community engagement and integration in health systems, but variable and uncertain levels of feasibility and acceptability of including a medical orientation to some elements of the curriculum through the inclusion of diagnostic and curative competencies

Annex 6.3 references

1. Bhutta Z, Memon Z, Soofi S, Salat M, Cousens S, Martines J. Implementing community-based perinatal care: results from a pilot study in rural Pakistan. Bulletin of the World Health Organization. 2008;6:452–9.

2. Bhutta ZA, Soofi S, Cousens S, Mohammad S, Memon ZA, Ali I et al. Improvement of perinatal and newborn care in rural Pakistan through community-based strategies: a cluster-randomised effectiveness trial. Lancet. 2011;377(9763):403–12.

6.4 *Modalities of pre-service training.* For CHWs receiving pre-service training, should the curriculum use specific delivery modalities versus not?

Summary of quantitative findings

Outcome(s)	Impact	No. of participants (studies)	Quality of evidence (GRADE)	Comments
Productivity: In-person training vs control (mailed resources)	One RCT reported two CHW productivity outcomes, with no effect reported for both outcomes (proportion of CHWs who delivered at least one brief intervention within past 30 days: mailed resources 78.7% vs in-person 75.6%, $P = 0.450$; within past 90 days: mailed resources 80.4% vs in-person 78.6%, $P = 0.649$)	547 (1)	⊕⊕ [a,b] Low	It is uncertain whether specific training modalities are more effective than others in improving CHW productivity
Productivity: Web-based training vs control (mailed resources)	One RCT reported two CHW productivity outcomes, with no effect reported for both outcomes (proportion of CHWs who delivered at least one brief intervention within past 30 days: mailed resources 78.7% vs web-based 76.0%, $P = 0.504$; within past 90 days: mailed resources 80.4% vs web-based 84.7%, $P = 0.147$)	547 (1)	⊕⊕ [a,b] Low	It is uncertain whether specific training modalities are more effective than others in improving CHW productivity
Knowledge: In-person training vs control (mailed resources)	One RCT reported two CHW knowledge outcomes, with positive effects reported for both outcomes, favouring the group receiving in-person training compared to a mailed resources group (tobacco and brief intervention knowledge – core: score difference 12.25, $P < 0.001$; tobacco and brief intervention knowledge – advanced: score difference 10.20, $P < 0.001$)	547 (1)	⊕⊕ [a,b] Low	It is uncertain whether specific training modalities are more effective than others in improving CHW knowledge
Knowledge: Web-based training vs control (mailed resources)	One RCT reported two CHW knowledge outcomes, with positive effects reported for both outcomes, favouring the group receiving web-based training compared to a mailed resources group (tobacco and brief intervention knowledge – core: score difference 14.61, $P < 0.001$; tobacco and brief intervention knowledge – advanced: score difference 14.82, $P < 0.001$)	547 (1)	⊕⊕ [a,b] Low	It is uncertain whether specific training modalities are more effective than others in improving CHW knowledge
Self-efficacy/esteem: In-person training vs control (mailed resources)	One RCT reported three CHW self-efficacy/esteem outcomes, with positive effects reported for all three outcomes, favouring the group receiving in-person training compared to a mailed resources group (confidence with brief intervention – score difference basic skills: 8.70, $P < 0.001$; motivational skills: 4.17, $P < 0.01$; quit planning: 4.54, $P < 0.01$)	547 (1)	⊕⊕ [a,b] Low	It is uncertain whether specific training modalities are more effective than others in improving CHW self-efficacy/esteem
Self-efficacy/esteem: Web-based training vs control (mailed resources)	One RCT reported three CHW self-efficacy/esteem outcomes, with positive effects reported for one outcome, favouring the group receiving in-person training compared to a mailed resources group, and no effect reported for two outcomes (confidence with brief intervention – score difference basic skills: 5.41, $P < 0.01$; motivational skills: 1.79, $P > 0.05$; quit planning: 0.66, $P > 0.05$)	547 (1)	⊕ [a,b,c] Very low	It is uncertain whether specific training modalities are more effective than others in improving CHW self-efficacy/esteem

Notes	Notes on GRADE scores
⊕ indicates that the overall quality of evidence is very low. ⊕⊕ indicates that the overall quality of evidence is low. ⊕⊕⊕ indicates that the overall quality of evidence is moderate. ⊕⊕⊕⊕ indicates that the overall quality of evidence is high.	a. Downgraded one level for risk of bias: information comes from study assessed as unclear or high risk of bias for the majority of domains. b. Downgraded one level for imprecision: event rate for dichotomous outcome < 300 or sample size for continuous outcome < 400. c. Downgraded one level for inconsistency: one outcome reported positive effect, two outcomes reported no effect.

GRADE quality assessment

Outcome(s)	Study design	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect (95% CI)	Dose–response gradient	Plausible control for confounding
Productivity: In-person training vs control (mailed resources)	RCT	Output	Direct	Delivery of tobacco brief interventions	547	1	6 months	●	No	No	Yes	No	No	No	No
Productivity: Web-based training vs control (mailed resources)	RCT	Output	Direct	Delivery of tobacco brief interventions	547	1	6 months	●	No	No	Yes	No	No	No	No
Knowledge: In-person training vs control (mailed resources)	RCT	Output	Indirect	Tobacco and brief intervention knowledge (core and advanced)	547	1	6 months	●	No	No	Yes	No	No	No	No
Knowledge: Web-based training vs control (mailed resources)	RCT	Output	Indirect	Tobacco and brief intervention knowledge (core and advanced)	547	1	6 months	●	No	No	Yes	No	No	No	No
Self-efficacy/esteem: In-person training vs control (mailed resources)	RCT	Output	Indirect	Confidence with brief intervention skills (basic; motivational; and quit plan)	547	1	6 months	●	No	No	Yes	No	No	No	No
Self-efficacy/esteem: Web-based training vs control (mailed resources)	RCT	Output	Indirect	Confidence with brief intervention skills (basic; motivational; and quit plan)	547	1	6 months	●	Yes	No	Yes	No	No	No	No

Legend

● Low risk of bias.
● Unclear risk of bias.
● High risk of bias.

Qualitative findings

Objective	To identify, appraise, and synthesize qualitative research evidence on the barriers and facilitators of the specific training modality of pre-service curriculum		
Perspective	Experiences and perspectives of CHWs on pre-service curriculum training modalities in any country		
Included programmes	CHW programmes		
Review finding	**Overall CERQual assessment of confidence**	**Explanation of CERQual assessment**	**studies contributing to the review**
Satisfaction: Classroom training was helpful	Moderate confidence	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Austin-Evelyn et al. *(1)*
Satisfaction: Flexibility of web-based training	Moderate confidence	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Castañeda et al. *(2)*
Self-efficacy/esteem: Classroom training increased self-worth, which contributed to well-being. Formal training also increased self-efficacy	Moderate confidence	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Castañeda et al. *(2)*, Morar et al. *(3)*
Morale: Training under the supervision of an experienced CHW and in clinical settings increased skills and morale	Moderate confidence	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Studies: Morar et al. *(3)*, Javanparast et al. *(4)*
Knowledge: Theory and practical classes increased knowledge	Moderate confidence	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Javanparast et al. *(4)*
Knowledge: Knowledge decreases over time and CHWs require refresher courses and supervision	Moderate confidence	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Austin-Evelyn et al. *(1)*
Advancement: In-person training allowed for the development of better interpersonal skills	Moderate confidence	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Studies: Castañeda et al. *(2)*, Morar et al. *(3)*
Advancement: Residential pre-service training influenced relationship building between students and trainers	Moderate confidence	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Javanparast et al. *(4)*
Studies: Austin-Evelyn et al. *(1)*, Castañeda et al. *(2)*, Morar et al. *(3)*, Javanparast et al. *(4)*			

Evidence to decision table

Recommendation 4

*WHO **suggests** using the following modalities for delivering pre-service training to CHWs:*

- balance of theory-focused knowledge and practice-focused skills, with priority emphasis on supervised practical experience;
- balance of face-to-face and e-learning, with priority emphasis on face-to-face learning, supplemented by e-learning on aspects on which it is relevant;
- prioritization of training in or near the community wherever possible;
- delivery of training and provision of learning materials in language that can optimize the trainees' acquisition of expertise and skills;
- ensuring a positive training environment;
- consideration of interprofessional training approaches where relevant and feasible.

Certainty of the evidence – very low. Strength of the recommendation – conditional.

Population: CHWs that undergo pre-service training

Intervention: use of specific modalities for delivery of pre-service training vs other modalities

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• The systematic review identified studies suggesting that training leads to indirect and developmental outputs for CHWs in the form of increased knowledge, advancement, self-efficacy/esteem, confidence and morale. While the studies described different training modalities, the evidence identified contains no clear indications of specific training modalities being more effective than others
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ○ Varies ● Don't know	• The systematic review did not identify any studies examining any harmful consequences related to different pre-service training delivery modalities
Certainty of evidence	●Very low ○ Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as very low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison	• The GDG was of the view that the most appropriate training modalities require finding the appropriate balance between different policy options

	○ Does not favour either the intervention or the comparison ○ Probably favours the intervention ○ Favours the intervention ● Varies ○ Don't know	
Resource use and cost-effectiveness	○ Large costs ○ Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ●Varies ○ Don't know	• The systematic review team did not identify any studies assessing cost-effectiveness of the policy options of interest. The policy options considered have different cost implications that should be examined when considering implementation • However, no cost-effectiveness evidence was found on this aspect
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ○ Probably increased ○ Increased ● Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. Some training modalities (e.g. training in communities/rural areas) are likely to be associated with improved equity
Acceptability and feasibility of intervention	○ No ○ Probably no ● Probably yes ○ Yes ○ Varies ○ Don't know	• Included qualitative studies indicate that different training modalities are acceptable by CHWs – the selection of modality may be context dependent. Decisions on which modality to use for training should consider trainee values, preferences and needs

Annex 6.4 references

1. Austin-Evelyn K, Rabkin M, Macheka T, Mutiti A, Mwansa-Kambafwile J, Dlamini T et al. Community health worker perspectives on a new primary health care initiative in the Eastern Cape of South Africa. PLoS ONE. 2017;12(3):e0173863.
2. Castañeda H, Nichter M, Nichter M, Muramoto M. Enabling and sustaining the activities of lay health influencers: lessons from a community-based tobacco cessation intervention study. Health Promotion Practice. 2010;11(4):483–92.
3. Morar NS, Naidoo S, Goolam A, Ramjee G. Research participants' skills development as HIV prevention peer educators in their communities. Journal of Health Psychology. 2016;1:1359105316655470.
4. Javanparast S, Baum F, Labonte R, Sanders D, Rajabi Z, Heidan G. The experience of community health workers training in Iran: a qualitative study. BMC Health Services Research. 2012:31(12):291.

6.5 *Competency-based certification.* For CHWs who have received pre-service training, should competency-based formal certification be used versus not?

Qualitative findings

Objective	To identify, appraise, and synthesize qualitative research evidence on the barriers and facilitators to the implementation of certification programmes for CHWs		
Perspective	Experiences and perspectives of CHWs on certification and credentialing in any country		
Included programmes	CHW programmes		
Review finding	**Overall CERQual assessment of confidence**	**Explanation of CERQual assessment**	**Studies contributing to the review**
Motivation: Trained and certified CHWs may earn higher wages	Moderate	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Kash, May and Tai-Seale *(1)*
Morale: Certification may enhance CHW recognition and acceptance at community and work level	Moderate	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Studies: Kash, May and Tai-Seale *(1)*, Amare *(2)*
Self-efficacy/esteem: Certification may improve CHW's self-esteem and self-worth	Moderate	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Kash, May and Tai-Seale *(1)*
Productivity: Trained and competent CHWs may provide better health care services	Moderate	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Kash, May and Tai-Seale *(1)*
Attrition: Certification may enhance CHW retention	Moderate	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Kash, May and Tai-Seale *(1)*
Advancement: Certification may promote career advancement and growth	Moderate	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Kash, May and Tai-Seale *(1)*
Feasibility: The implementation of credentialing or certification processes may face a range of implementation challenges	Moderate	Minor concerns regarding methodological limitations, relevance, coherence, and adequacy	Study: Catalani et al. *(3)*

Evidence to decision table

Recommendation 5 *WHO **suggests** using competency-based formal certification for CHWs who have successfully completed pre-service training.* Certainty of the evidence – very low. Strength of the recommendation – conditional.		
Population: CHWs that undergo pre-service training Intervention: competency-based formal certification		
Factors	**Decision**	**Explanations/comments**
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• The qualitative evidence included in these studies points to potential, but untested, benefits from certification processes related to CHW motivation, morale and self-esteem, as well as their retention, professional development and advancement • The process of credentialing was perceived by CHWs as offering opportunities to gain increased knowledge, credibility and recognition
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ○ Varies ● Don't know	• The literature also points to possible barriers to the implementation of credentialing in that certification requirements may impose costs and resource demands on CHWs, and limit the accessibility of community health service positions to volunteers who are interested in working in the sector but not eligible or suited for certification
Certainty of evidence	● Very low ○ Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as very low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ○ Probably favours the intervention ○ Favours the intervention ○ Varies ● Don't know	• Despite the paucity of evidence, the GDG concluded that the potential benefits of certification in terms of institutionalization of CHWs and career progression opportunities might outweigh potential unintended effects

Resource use and cost-effectiveness	○ Large costs ● Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• The moderate costs likely to be required for accreditation are justified in relation to the potential benefits of this policy option • However, no cost-effectiveness evidence was found on this aspect
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ○ Probably increased ○ Increased ● Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The GDG recognized the need to deliberately design the accreditation process to include equity considerations
Acceptability and feasibility of intervention	○ No ○ Probably no ● Probably yes ○ Yes ○ Varies ○ Don't know	• The systematic review team identified studies whose participants generally viewed credentialing as positive, but one study also pointed to potential implementation challenges • The stakeholder perception survey found competency-based certification to be both acceptable and feasible

Annex 6.5 references

1. Kash BA, May ML, Tai-Seale M. Community health worker training and certification programs in the United States: findings from a national survey. Health Policy (Amsterdam, Netherlands). 2007;80(1):32–42.
2. Amare Y. Non-financial incentives for voluntary community health workers: a qualitative study. Working Paper No. 2. The Last Ten Kilometers Project. Addis Ababa, Ethiopia: Training Institute, Inc.; 2011.
3. Catalani CE, Findley SE, Matos S, Rodriguez R. Community health worker insights on their training and certification. Progress in Community Health Partnerships. 2009;3(3):201–2.

6.6 *Supportive supervision.* In the context of CHW programmes, what strategies of supportive supervision should be adopted over what other strategies?

Quantitative findings: overview of included quantitative studies

Study	Design	Setting and participants	Intervention	Comparison or control	Measures, data collection	Findings	Outcomes
Intervention studies							
Chang et al. *(1)*	RCT	29 community-based peer health workers	Health sciences programme staff member provided day-to-day supervision of the CHWs. CHWs working at 10 clinics were randomized by clinic to receive the intervention (phones to call and text higher-level providers with patient-specific clinical information)	No mHealth intervention	Patients' risk of virological failure	There were no significant differences in patients' risk of virological failure (RR 1.17; 95% CI, 0.84–1.64, $P = 0.34$)	Impact level
Das et al. *(2)*	RCT	120 villages in two districts of Odisha, India	Supportive supervision of ASHAs was combined with community mobilization in intervention arm A. Arm B was provided with only community mobilization activities	Observational control arm	Outcome measures included changes in the utilization of bednets and timely care seeking for fever from a trained provider compared to the control group	Significantly more respondents slept under a bednet the previous night in arm A (84.54%; 95% CI, 1.328–1.661, $P < 0.001$) and arm B (82.43%; 95% CI, 1.143–1.419, $P < 0.001$) than in the control arm (78.65%), and fever incidence in treatment villages was lower than in the control villages	Service delivery/impact level
Singh et al. *(3)*	RCT	CHWs and CHVs in eight villages in Budondo subcounty in Jinja, Uganda	CHWs trained and supportively supervised community health volunteers (CHVs) in Uganda to deliver education about pregnancy, newborn care, family planning and hygiene	No supervision	The study compared training alone vs training and supportive supervision by paid CHWs on the effectiveness of CHVs to deliver education about pregnancy, newborn care, family planning and hygiene	The study showed that 95% of all CHVs were retained Increased numbers of home visits Increased community participation measured by the increase in the number of community-built handwashing devices	Service delivery/process level

Study	Design	Setting and participants	Intervention	Comparison or control	Measures, data collection	Findings	Outcomes
Som et al. *(4)*	Quasi-experimental	Six districts in India	The study included all supervisors who had undergone supportive supervision training and were posted in the intervention district	No supportive supervision training	The study measured knowledge amongst CHWs on the correct maintenance of a temperature logbook, cabinet temperature, placing of diluents and manner of placing ice packs	The knowledge of CHWs receiving training was better than those who did not	Service delivery/process level
Ayele, Desta and Larson *(5)*	RCT	Community health agents (CHAs) located in two districts of the south-western Ethiopian region of Illubabor	CHAs received a refresher course and monthly supervision	Those who did not receive any intervention	Assessed functional status (presence of any one of 13 criteria related to specifications in CHA job description)	Refresher training and regular supervision had a significant impact on the functional status of CHWs The mean composite functional status score for the intervention group increased from 13.1 at onset to 18.4 at 2 months follow-up and thereafter remained stable with a gradual increase to 19.26 at 6 months	Service delivery level
Kaphle, Matheke-Fischer and Lesh *(6)*	RCT	60 CHWs in five districts of Madhya Pradesh, India	CHWs received performance feedback and supportive supervision	No feedback and supportive supervision	Assessment on three performance indicators: case activity (number of clients visited), form submissions, and duration of counselling	There was a significant impact of supervision on duration of counselling (increased by 3.86 minutes, $P < 0.004$) The number of form submissions did not show significant improvements	Service delivery level
DeRenzi et al. *(7)*	RCT	CHWs in United Republic of Tanzania	Automated SMS reminder system on CHW performance with regard to CHW visit responsiveness	Did not receive reminders	CHW responsiveness was measured by frequency matching using baseline performance data	The reminders resulted in an 86% reduction in the average number of days a CHW's clients were overdue (9.7 to 1.4 days), with only a small number of cases ever escalating to the supervisor	Service delivery level

Study	Design	Setting and participants	Intervention	Comparison or control	Measures, data collection	Findings	Outcomes
Observational studies							
Ameha et al. *(8)*	Descriptive	Health extension workers (HEWs) in 113 districts in Ethiopia	NA	NA	Assessed the effectiveness of supervisory visits to improve the consistency of integrated community case management (iCCM) skills through descriptive analysis of records	Identified a positive dose–response relationship between the number of supervisory visits and iCCM skills of HEWs	Service delivery level
Nonaka et al. *(9)*	Before–after study design	154 villages located in Xepone district, Lao People's Democratic Republic	Phone distribution and its impact on event surveillance reports	NA	Assess the usefulness of a mobile phone-based communication network between village health volunteers (VHVs) and their supervisors and determined productivity of the VHVs	Submissions of monthly vital event surveillance reports significantly increased from 79 (51.3%) to 127 (82.5%) at 6 months after phone distribution An increase was also noted in treatment consultations, material requesting and meeting scheduling. This increase was maintained at one year follow-up Calls with supervisors addressed reporting, advice seeking, medical supply requests, and scheduling of outreach activities	Service delivery level
Rowe et al. *(10)*	Cross-sectional study design	114 CHWs in Siaya district, Kenya	The study tested two models: model 1 compared treatment with no error in managing childhood illness vs treatment with a minor or major error; and model 2 compared treatment with a major error vs treatment with no error or a minor error		Analysed to assess the effect of intervention-related factors: refresher training, supervision, involvement of community women in the CHW selection process, adequacy of medicine supplies, and use of a guideline flipchart during consultations	The number of supervisory contacts was not significantly associated with adherence to clinical guidelines in both models of the study (model 1 treatment with no error: adjusted OR 1.04; 95% CI, 0.92–1.18, $P = 0.52$, vs model 2 treatment with major error: adjusted OR 1.02; 95% CI, 0.92–1.12, $P = 0.74$)	

Study	Design	Setting and participants	Intervention	Comparison or control	Measures, data collection	Findings	Outcomes
Definitions **Odds ratio (OR).** A measure of effect that is used to approximate relative risk (i.e., the likelihood that one group will experience the outcome given a certain exposure versus the likelihood that another group will experience the outcome given they were not exposed). When the OR is greater than 1.0, the risk is greater. When the OR is between 0 and 1, the risk is lower. When the risk is 1.0, there is no difference between groups. The further the OR is above or below 1.0, the larger the effect. **Correlation.** A measure of association between two different constructs. **Significance or statistical significance.** The probability that a finding was observed by chance alone. Traditionally, a finding is said to be "significant" when this probability is less than 0.05 (i.e., $P < 0.05$). **Confidence interval (CI).** The estimated interval between which the measure of effect (e.g. the OR) would probably be observed if the study were conducted again on a similar sample of subjects. **Adjusted (e.g. adjusted OR).** When an explanatory or causal factor's raw association with an outcome is statistically adjusted to take account of other potential explanatory factors.							

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Quality of evidence (GRADE)	Comments
	Control risk	Intervention risk			
Epidemic control: treatment vs control	NA	NA	0.23 (0.15–0.31)	⨁ [a]	
Slept last night under a bednet	NA	NA	1.49 (1.32–1.66)	⨁⨁⨁	There is moderate evidence from one RCT that supportive supervision of CHW can affect health outcomes at a community level
Bednet ownership: households with at least one bednet	NA	NA	0.63 (0.21–1.95)	⨁⨁ [b]	It is uncertain whether supportive supervision of CHW can affect health outcomes at a community level
Fever diagnosed by a CHW	NA	NA	1.642 (1.16–2.312)	⨁⨁⨁ [b]	There is moderate evidence from one RCT that supportive supervision of CHW can affect health outcomes at a community level
Patients' risk of virological failure associated with the use of mHealth			1.17 (0.84–1.64)	⨁⨁	It is uncertain whether supportive supervision of CHW can affect health outcomes at a community level

All effects are intervention vs control.

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Notes on GRADE scores

a. Cohen's d (standardized difference in means between treatment and control).

b. Odds ratio: supportive supervision and community mobilization versus control. Note: *n* represents number of households.

GRADE quality assessment

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Other considerations	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)
Health care promoting behaviour																
Slept last night under a bednet (RCT)	Outcome	CHW-attributable changes among individual clients	Health care behaviour promoting	2344	1: Das et al *(2)*		🟡	No [1]	No	No	No [a]	No	No	No	No	⨁⨁⨁
Bednet ownership: households with at least one bednet (RCT)	Outcome	CHW-attributable changes among individual clients	Health care behaviour promoting	2344	1: Das et al *(2)*		🟡	No [1]	No	Yes	No [a]	No	No	No	No	⨁⨁, [b,c]
Health care seeking behaviour																
Fever diagnosed by a CHW (RCT)	Outcome	CHW-attributable changes among individual clients	Health care seeking behaviour	2344	1: Das et al *(2)*		🟡	No [1]	No	No	No [a]	No	No	No	No	⨁⨁⨁, [b]
Impact																
Patients' risk of virological failure associated with the use of mHealth (RCT)	Outcome	Direct	Impact	970	1: Chang et al. *(1)*	2 years	🔴	No [1]	No	Yes	No [a]	No	No	No	No	⨁, [b,c]

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Notes on GRADE scores

a. Only one study for the outcome; thus cannot be assessed.

b. Downgraded one level for imprecision: wide confidence interval.

c. Downgraded one level for risk of bias: information comes from studies assessed as high risk or unclear of bias for the majority of domains.

Legend

🟢 Low risk of bias.

🟡 Unclear risk of bias.

🔴 High risk of bias.

Newcastle-Ottawa quality assessment for cross-sectional studies

Study	Selection				Comparability	Outcome		Total score
	Representativeness of sample	Sample size	Non-respondents	Ascertainment of exposure		Assessment of outcome	Appropriateness of statistical test	
Possible score	1	1	1	2	2	1	2	Max: 10
Rowe et al. *(10)*	●	●		●	●●	●	●	7
Ameha et al. *(8)*	●	●		●	●	●	●	6

Qualitative findings

Objective	To identify supportive supervision mechanisms for community health workers		
Perspective	Opinion of stakeholders on supportive supervision		
Included programmes	Programme of CHWs in a low- and middle-income country		
Review finding	**Overall CERQual assessment of confidence**	**Explanation of CERQual assessment**	**Studies contributing to the review**
Role of supervisors is important, both professionally and emotionally	High confidence	This evidence is graded as high confidence as evidence is from four studies with minimal concern on all indicators	Chang et al. *(1)*, Henry et al. *(11)*, Rabbani et al. *(12)*, Callaghan-Koru et al. *(13)*
The presence of supervisors during visit improves performance of CHWs	Low confidence	This evidence is graded as low confidence as evidence is from a single study with moderate concern on methodology and adequacy	Rabbani et al. *(12)*
CHWs preferred verbal group feedback while supervisors preferred individual written feedback	Low confidence	This evidence is graded as low confidence as evidence is from a single study with moderate concern on methodology and adequacy	Rabbani et al. *(12)*
With minimal training, CHWs and their supervisors tailored the multimodal communication features of the mobile to enact virtual one-to-one, group, and peer-to-peer forms of supervision and support	Low confidence	This evidence is graded as low confidence as it is from one study that explores aspects of communication of WhatsApp tool. The results are coherent, but the limited number of participants and use in a setting where the particular tool is very common explores questions on generalizability of the data	Henry et al. *(11)*

Evidence to decision table

Recommendation 6

*WHO **suggests** using the following supportive supervision strategies in the context of CHW programmes:*

- appropriate supervisor–supervisee ratio allowing meaningful and regular support;
- ensuring supervisors receive adequate training;
- coaching and mentoring of CHWs;
- use of observation of service delivery, performance data and community feedback;
- prioritization of improving the quality of supervision.

Certainty of the evidence – very low. Strength of the recommendation – conditional.

Population: practising CHWs

Intervention: supportive supervision strategies vs other strategies

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ● Large ○ Varies ○ Don't know	• The systematic review identified quantitative and qualitative studies suggesting that supervision is a strategy that may enhance the quality of the work of CHWs. The studies, however, did not provide specific evidence on individual supervision approaches and strategies
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• The systematic review did not find any studies examining any harmful or unintended consequences of supportive supervision
Certainty of evidence	● Very low ○ Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as very low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the	• The GDG considered that the overwhelming evidence on the positive results of supportive supervision and the lack of known or theoretical harms probably favours the adoption of

	comparison ● Probably favours the intervention ○ Favours the intervention ○ Varies ○ Don't know	different supervision strategies, notwithstanding the uncertainty of the comparative evidence on different supervision strategies
Resource use and cost-effectiveness	○ Large costs ● Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• The moderate costs likely to be required for supportive supervision are justified, given its contribution to improved quality and motivation of CHWs • However, no cost-effectiveness evidence was found on this aspect
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The GDG was of the view that the adoption of different supportive supervision strategies is likely to increase health equity through improved capacity and performance of CHWs in delivering health care to underserved communities
Acceptability and feasibility of intervention	○ No ○ Probably no ● Probably yes ○ Yes ○ Varies ○ Don't know	• The systematic review team identified studies that, while not focused on feasibility, suggest the general feasibility of supervision in the context of CHW programmes • The stakeholder perception survey identified most supportive supervision strategies to be acceptable and feasible, but lower levels of acceptability and especially of feasibility of direct supervision of service delivery and of supervision conducted by other CHWs

Annex 6.6 references

1. Chang LW, Kagaayi J, Arem H, Nakigozi G, Ssempijja V, Serwadda D 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):1776–84.
2. Das A, Friedman J, Kandpal E, Ramana GN, Gupta RK, Pradhan MM et al. Strengthening malaria service delivery through supportive supervision and community mobilization in an endemic Indian setting: an evaluation of nested delivery models. Malaria Journal. 2014;13:482.
3. Singh D, Negin J, Orach CG, Cumming R. Supportive supervision for volunteers to deliver reproductive health education: a cluster randomized trial. Reproductive Health. 2016;13(1):126.
4. Som M, Panda B, Pati S, Nallala S, Anasuya A, Chauhan AS et al. Effect of supportive supervision on routine immunization service delivery: a randomized post-test study in Odisha. Global Journal of Health Science. 2014;6(6):61–7.

5. Ayele F, Desta A, Larson C. The functional status of community health agents: a trial of refresher courses and regular supervision. Health Policy and Planning. 1993;8(4):379–84.

6. Kaphle S, Matheke-Fischer M, Lesh N. Effect of performance feedback on community health workers motivation and performance in Madhya Pradesh, India: a randomized controlled trial. JMIR Public Health and Surveillance. 2016;2(2):e169.

7. DeRenzi B, Findlater L, Payne J, Birnbaum B, Mangilima J, Parikh T et al. Improving community health worker performance through automated SMS. 2012.

8. Ameha A, Karim AM, Erbo A, Ashenafi A, Hailu M, Hailu B et al. Effectiveness of supportive supervision on the consistency of integrated community cases management skills of the health extension workers in 113 districts of Ethiopia. Ethiopian Medical Journal. 2014;52(Suppl. 3):65–71.

9. Nonaka D, Pongvongsa T, Nishimoto F, Nansounthavong P, Hongwei J, Vongsouvanh A et al. Successful mobile phone network-based approach to integration of the health care system in rural Laos: strengthening lay health worker performance. Rural and Remote Health. 2014;14:2588.

10. Rowe SY, Kelly JM, Olewe MA, Kleinbaum DG, McGowan JE Jr, McFarland DA et al. Effect of multiple interventions on community health workers' adherence to clinical guidelines in Siaya district, Kenya. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2007;101(2):188–202.

11. Henry JV, Winters N, Lakati A, Oliver M, Geniets A, Mbae SM et al. Enhancing the supervision of community health workers with WhatsApp mobile messaging: qualitative findings from 2 low-resource settings in Kenya. Global Health Science and Practice. 2016;4(2):311–25.

12. Rabbani F, Perveen S, Aftab W, Zahidie A, Sangrasi K, Qazi SA. Health workers' perspectives, knowledge and skills regarding community case management of childhood diarrhoea and pneumonia: a qualitative inquiry for an implementation research project "Nigraan" in district Badin, Sindh, Pakistan. BMC Health Services Research. 2016;16:462.

13. Callaghan-Koru JA, Hyder AA, George A, Gilroy KE, Nsona H, Mtimuni A et al. Health workers' and managers' perceptions of the integrated community case management program for childhood illness in Malawi: the importance of expanding access to child health services. American Journal of Tropical Medicine and Hygiene. 2012;87(Suppl. 5):61–8.

6.7 *Remuneration.* In the context of CHW programmes, should practising CHWs be paid for their work versus not?

Quantitative findings: overview of included quantitative studies

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
Intervention studies							
Bossuroy, Delavallade and Pons *(1)*	RCT	Assessing the impact of an enhanced directly observed treatment, short course (DOTS) model in community	CHWs were offered performance incentives in addition to their base salary	Crossover. In the first 6 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 6 months, they were randomly reassigned to either a fixed or an incentivized salary scheme based on the treatment default rates	In the first 6 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 6 months, they were randomly reassigned to either a fixed or an incentivized salary scheme based on the treatment default rates. These incentives (for detection) or penalties (for default) were added to a base salary that guaranteed the health workers a minimum income	The incentives increased the number of new detections of tuberculosis (TB) by 24.1% each month With the inclusion of city-fixed effects (the level of stratification) and health worker control variables, the number of new detections of TB increased to 33.2% ($P < 0.05$) The incentives were also noted to decrease CHW job satisfaction by about 0.25 standard deviations	Knowledge, competency, service delivery
Andreoni et al. *(2)*	RCT	Door-to-door immunization services	A base salary of US$ 1 daily with a fixed bonus of US$ 10	Performance of CHWs with "phone only" were compared with CHWs with "phone plus incentive"	Assessed CHW performance in two vaccination drives	Results show that overall productivity increases when tailored contracts consider performance preferences as well as time preferences of CHWs (standardized mean difference (SMD) –0.32; –0.54 to –0.1)	Service delivery
Observational studies							
Adejumo et al. *(3)*	Cross-sectional study	Four different active case-finding strategies for boosting TB case detection in three Nigerian states	NA	NA	These models varied in their mode of recruitment, frequency, method of supervision and motivation of workers, record keeping, level and regularity of compensation	The study found that incentivized referral, appropriate selection of CHWs, supportive supervision, leveraged treatment support roles and a responsive TB program to receive clients for testing were the key drivers of community TB case finding	Service delivery

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
Srivastava et al. *(4)*	Cross-sectional study	Accredited social health activists (ASHAs) in Uttar Pradesh, India	NA	NA	Assessed training, functioning, knowledge and acceptability. Data were collected through predesigned and pretested structured interviews	Major motivating factors for ASHAs were either money (81.66%) or incorporation into a government job (66.66%). ASHAs were also interested in charity (43.33%), and a good number of them also put in extra efforts to gain their financial incentives (25%)	Motivation and service delivery
Rowe et al. *(5)*	Cross-sectional study	CHW adherence to clinical guidelines in Kenya	NA	NA	Adherence to clinical guidelines measured and compared against incentives provided	Adherence scores were significantly higher for CHWs who thought that they received four or five benefits while working, including earning money. If the CHW had another job earning money, however, there was no difference in the overall weighted adherence to guidelines compared with CHWs who did not have another job	Motivation and service delivery

Definitions

Odds ratio (OR). A measure of effect that is used to approximate relative risk (i.e., the likelihood that one group will experience the outcome given a certain exposure versus the likelihood that another group will experience the outcome given they were not exposed). When the OR is greater than 1.0, the risk is greater. When the OR is between 0 and 1, the risk is lower. When the risk is 1.0, there is no difference between groups. The further the OR is above or below 1.0, the larger the effect.

Correlation. A measure of association between two different constructs.

Significance or statistical significance. The probability that a finding was observed by chance alone. Traditionally, a finding is said to be "significant" when this probability is less than 0.05 (i.e., $P < 0.05$).

Confidence interval (CI). The estimated interval between which the measure of effect (e.g. the OR) would probably be observed if the study were conducted again on a similar sample of subjects.

Adjusted (e.g. adjusted OR). When an explanatory or causal factor's raw association with an outcome is statistically adjusted to take account of other potential explanatory factors.

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
Case detection-based incentives	NA	TB case detection per month: 24.1% increase	Not possible to determine	2760 (1)	⨁⨁ [a]
Comprehensive quota-oriented model (training, supervision and US$ 80 quarterly payment)	NA	Median referrals: 48 per CHW per year Mean TB diagnoses: 7.1 per CHW per year	Not possible to determine	115 (1)	⨁⨁ [a]
Vaccination distribution according to policy goals via tailored contracts	NA	NA	SMD –0.32; –0.54 to –0.1	349 (1)	⨁⨁ [b]
Identification of major motivating factors for ASHAs	NA	81.66% of respondents reported that money is a motivating factor	Not possible to determine	226 (1)	⨁⨁ [c]
Adherence to clinical guidelines	NA	CHWs that thought they received four or five benefits achieved clinical guideline adherence in 82.5% of cases	Not possible to determine	114 (1)	⨁⨁ [c]

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Summary findings

a. There is limited evidence that payment of CHWs can lead to changes in service delivery.

b. There is limited evidence that payment of CHWs can lead to CHW-attributable changes among individual clients.

c. There is limited evidence that payment of CHWs can lead to CHW-attributable changes among individual clients.

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)	Comments
	Control risk	Intervention risk				
Case detection-based incentives	NA	TB case detection per month: 24.1% increase	Not possible to determine	2760 (1)	⊕⊕	There is limited evidence that payment of CHW can lead to changes in service delivery
Comprehensive quota-oriented model (training, supervision and US$ 80 quarterly payment)	NA	Median referrals: 48 per CHW per year Mean TB diagnoses: 7.1 per CHW per year	Not possible to determine	115 (1)	⊕	There is limited evidence that payment of CHW can lead to changes in service delivery
Vaccination distribution according to policy goals via tailored contracts	NA	NA	SMD –0.32; –0.54 to –0.1	349 (1)	⊕⊕	There limited evidence from one RCT that payment of CHW can lead to CHW-attributable changes among individual clients
Identification of major motivating factors for ASHAs	NA	81.66% of respondents reported that money is a motivating factor	Not possible to determine	226 (1)	⊕	There is limited evidence that payment of CHW can lead to CHW-attributable changes among individual clients
Adherence to clinical guidelines	NA	CHWs that thought they received four or five benefits achieved clinical guideline adherence in 82.5% of cases	Not possible to determine	114 (1)	⊕	There is limited evidence that payment of CHW can lead to CHW-attributable changes among individual clients

Notes

⊕ indicates that the overall quality of evidence is very low.

⊕⊕ indicates that the overall quality of evidence is low.

⊕⊕⊕ indicates that the overall quality of evidence is moderate.

⊕⊕⊕⊕ indicates that the overall quality of evidence is high.

GRADE quality assessment

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Other considerations	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)
Case detection-based incentives (RCT)	Outcome	CHW-attributable changes among individual clients	Service delivery	2760	1: Bossuroy, Delavallade and Pons *(1)*	NA	🔴	No	No	No	No [a]	No	No	No	No	⨁⨁ [b]
Vaccination distribution according to policy goals via tailored contracts (RCT)	Output	Direct	Service delivery	349	1: Andreoni et al. *(2)*	NA	🟡	No	No	Yes	No [a]	No	No	No	No	⨁⨁ [c,d]
Comprehensive quota-oriented model (training, supervision and US$ 80 quarterly payment) (non-RCT)	Outcome	CHW-attributable changes among individual clients	Service delivery	115	1: Adejumo et al. *(3)*	NA	🟡	No	No	Yes	No [a]	No	No	No	No	⨁ [c,d]
Identification of major motivating factors for ASHAs (non-RCT)	Outcome	CHW-attributable changes among individual clients	Motivation	226	1: Srivastava et al. *(4)*	NA	🟡	No	No	Yes	No [a]	No	No	No	No	⨁ [c]
Adherence to clinical guidelines (non-RCT)	Outcome	CHW-attributable changes among individual clients	Motivation	114	1: Rowe et al. *(5)*	NA	🟡	No	No	Yes	No [a]	No	No	No	No	⨁ [c]

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Notes on GRADE scores

a. Only one study for the outcome; thus cannot be assessed.

b. Downgraded one level for imprecision: wide confidence interval.

Legend

🟢 Low risk of bias.

🟡 Unclear risk of bias.

🔴 High risk of bias.

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Other considerations	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)
					c. Downgraded one level for risk of bias: information comes from studies assessed as high risk or unclear of bias for the majority of domains.											

Newcastle-Ottawa quality assessment for cross-sectional studies

Study	Selection				Comparability	Outcome		Total score
	Representativeness of sample	Sample size	Non-respondents	Ascertainment of exposure		Assessment of outcome	Appropriateness of statistical test	
Possible score	1	1	1	2	2	1	2	Max: 10
Adejumo et al. *(3)*	✹	✹		✹	✹	✹	✹	6
Srivastava et al. *(4)*	✹	✹		✹	✹	✹	✹	6
Rowe et al. *(5)*				✹	✹	✹	✹	4

Qualitative findings

Objective	To identify incentives for community health workers		
Perspective	Experience among stakeholders		
Included programmes	Two programmes on performance-based incentives and one on financial incentives		
Review finding	**Overall CERQual assessment of confidence**	**Explanation of CERQual assessment**	**Studies contributing to the review**
Financial incentives are perceived as important for the motivation and performance of CHWs	High confidence	Minor concerns in methodology, relevance, coherence and adequacy	Maes and Kalofonos *(6)*, Schwarz et al. *(7)*, Scott and Shanker *(8)*, Dil et al. *(9)*, Takasugi and Lee *(10)*, Sarin et al. *(11)*, Condo et al. *(12)*, Miller et al. *(13)*, Singh et al. *(14)*
Performance-based incentive models are solely based on end results and do not acknowledge the total effort	Moderate confidence	Moderate concerns in methodology, adequacy and relevance	Scott and Shanker *(8)*, Sarin et al. *(11)*, Miller et al. *(13)*
Performance-based incentive models discouraged CHWs who earned less	Low confidence	Moderate concerns in methodology, adequacy and relevance	Miller et al. *(13)*
Getting paid reduced the trust among community members, as they believed CHWs were working for their personal gains	Low confidence	Moderate concerns in methodology, adequacy and relevance	Sarin et al. *(11)*, Miller et al. *(13)*
Performance-based incentive models inspired a sense of financial independence and self-confidence	Low confidence	Moderate concerns in methodology, adequacy and relevance	Sarin et al. *(11)*
Mixed salary better than incentive and performance-based incentive for some	Low confidence	Moderate concerns in methodology, adequacy and relevance	Sarin et al. *(11)*, Miller et al. *(13)*
Certain activities may have been prioritized over others, due to being linked to higher incentives	Low confidence	Moderate concerns in methodology, adequacy and relevance	Sarin et al. *(11)*, Miller et al. *(13)*

Evidence to decision table

Recommendations

*7A: WHO **recommends** remunerating practising CHWs for their work with a financial package commensurate with the job demands, complexity, number of hours, training and roles that they undertake.*

Certainty of the evidence – very low. Strength of the recommendation – strong.

*7B: WHO **suggests** not paying CHWs exclusively or predominantly according to performance-based incentives.*

Certainty of the evidence – very low. Strength of the recommendation – conditional.

Population: practising CHWs

Intervention: providing a financial package to practising CHWs vs not

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ● Large ○ Varies ○ Don't know	• The systematic review team identified a few quantitative studies showing that financial incentives may lead to improved CHW performance • Qualitative studies were included to provide insights with respect to the perceived consequences of various payment and incentivization approaches. Financial incentives in general appear to be well accepted, provide motivation, and may bring a sense of financial independence and self-confidence to CHWs • The systematic review of reviews identified evidence that monetary remuneration and non-monetary incentives are important motivators for CHWs
Magnitude of undesirable effects	○ Trivial ● Small ○ Moderate ○ Large ○ Varies ○ Don't know	• Negative results and perceptions were primarily linked to performance-based incentive schemes. They were described as at times being too narrowly focused on pre-identified indicators, leading to activities and efforts not linked to these indicators being ignored and unacknowledged
Certainty of evidence	● Very low ○ Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the	• The GDG was of the view that the evidence overwhelmingly favours the intervention of providing CHWs with a financial package

	comparison ○ Probably favours the intervention ● Favours the intervention ○ Varies ○ Don't know	• However, with regard to the provision of performance-based incentives, the GDG was of the view that the balance varies and might be in favour of the comparison (i.e. not paying performance-based incentives)
Resource use and cost-effectiveness	● Large costs ○ Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• The GDG recognized that the policy of providing a financial package to practising CHWs has major cost implications that should be factored into human resources for health and health system planning through the appropriate mobilization and allocation of adequate financial resources • No cost-effectiveness evidence was found on CHW remuneration
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The GDG was of the view that providing CHWs with a financial package is likely to increase health equity through improved motivation and performance of CHWs in delivering health care to underserved communities • The GDG was however concerned that performance-based incentive schemes could lead to skewed priorities, with potentially negative impacts on health equity
Acceptability and feasibility of intervention	○ No ○ Probably no ○ Probably yes ○ Yes ● Varies ○ Don't know	• The systematic review team identified evidence that standard financial incentives were perceived as acceptable, motivating, providing a sense of financial independence and self-confidence amongst community workers and improving their performance. However, unstable payment systems creating delays or uneven payment led to dissatisfaction among CHWs. Performance-based incentive schemes raised concerns as they direct activities towards indicators and potentially lead to other activities being ignored • The stakeholder perception survey identified a good level of feasibility and acceptability of providing CHWs with both financial and non-financial incentives, but the assessment of feasibility of paying CHWs a minimum wage bordered on the uncertainty range • The GDG was of the view that the acceptability and feasibility of the policy option would vary considerably according to the country context and health policy priorities

Annex 6.7 references

1. Bossuroy T, Delavallade C, Pons V. Fighting tuberculosis through community based counsellors: a randomized evaluation of performance based incentives in India. New Delhi: International Initiative for Impact Evaluation; 2016.

2. Andreoni J, Callen M, Khan Y, Jaffar K, Sprenger C. Using preference estimates to customize incentives: an application to polio vaccination drives in Pakistan. National Bureau of Economic Research; 2016.

3. Adejumo AO, Azuogu B, Okorie O, Lawal OM, Onazi OJ, Gidado M et al. Community referral for presumptive TB in Nigeria: a comparison of four models of active case finding. BMC Public Health. 2016;16:177.

4. Srivastava D, Prakash S, Adhish V, Nair K, Gupta S, Nandan D. A study of interface of ASHA with the community and the service providers in eastern Uttar Pradesh. Indian Journal of Public Health. 2009;53(3):133–6.

5. Rowe SY, Kelly JM, Olewe MA, Kleinbaum DG, McGowan JE Jr, McFarland DA et al. Effect of multiple interventions on community health workers' adherence to clinical guidelines in Siaya district, Kenya. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2007;101(2):188–202.

6. Maes K, Kalofonos I. Becoming and remaining community health workers: perspectives from Ethiopia and Mozambique. Social Science and Medicine. 2013;87:52–9.

7. Schwarz D, Sharma R, Bashyal C, Schwarz R, Baruwal A, Karelas G et al. Strengthening Nepal's female community health volunteer network: a qualitative study of experiences at two years. BMC Health Services Research. 2014;14:473.

8. Scott K, Shanker S. Tying their hands? Institutional obstacles to the success of the ASHA community health worker programme in rural north India. AIDS Care. 2010;22(Suppl. 2):1606–12.

9. Dil Y, Strachan D, Cairncross S, Korkor AS, Hill Z. Motivations and challenges of community-based surveillance volunteers in the Northern region of Ghana. Journal of Community Health. 2012;37(6):1192–8.

10. Takasugi T, Lee A. Why do community health workers volunteer? A qualitative study in Kenya. Public Health. 2012;126(10):839–45.

11. Sarin E, Lunsford SS, Sooden A, Rai S, Livesley N. The mixed nature of incentives for community health workers: lessons from a qualitative study in two districts in India. Frontiers in Public Health. 2016;4:38.

12. Condo J, Mugeni C, Naughton B, Hall K, Tuazon MA, Omwega A et al. Rwanda's evolving community health worker system: a qualitative assessment of client and provider perspectives. Human Resources for Health. 2014;12:71.

13. Miller JS, Musominali S, Baganizi M, Paccione GA. A process evaluation of performance-based incentives for village health workers in Kisoro district, Uganda. Human Resources for Health. 2014;12:19.

14. Singh D, Negin J, Otim M, Orach CG, Cumming R. The effect of payment and incentives on motivation and focus of community health workers: five case studies from low- and middle-income countries. Human Resources for Health. 2015;13:58.

6.8 *Contracting agreements.* In the context of CHW programmes, should practising CHWs have a formal contract versus not?

Quantitative findings: overview of included quantitative studies

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)	Comments
	Control risk	Intervention risk				
Performance of CHW following receipt of an appointment letter	0.156	0.395	3.57; 2.00–6.26	1	⊕, [a]	There is limited evidence that the presence of a formal contract has any impact on service delivery
Closeness to vaccine distribution policy objectives with use of tailored contract			SMD –0.32; –0.54 to –0.1	1	⊕⊕, [a]	There is limited evidence that the presence of a formal contract has any impact on service delivery

Notes

⊕ indicates that the overall quality of evidence is very low.

⊕⊕ indicates that the overall quality of evidence is low.

⊕⊕⊕ indicates that the overall quality of evidence is moderate.

⊕⊕⊕⊕ indicates that the overall quality of evidence is high.

GRADE quality assessment

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Publication bias	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)
Closeness to vaccine distribution policy objectives with use of tailored contract (RCT)	Output	Direct	Service delivery	349	1	NA	●	No	No	Yes	No	No	No	No	No	⊕⊕ [b,c]
Performance of CHW following receipt of an appointment letter (non-RCT)	Outputs	Direct	Service delivery	336	1	NA	●	No [a]	No	Yes	No	No	No	No	No	⊕ [c]

Notes

⊕ indicates that the overall quality of evidence is very low.

⊕⊕ indicates that the overall quality of evidence is low.

⊕⊕⊕ indicates that the overall quality of evidence is moderate.

⊕⊕⊕⊕ indicates that the overall quality of evidence is high.

Notes on GRADE scores

a. Not possible to assess as outcome is for one study only.

b. Downgraded one level for risk of bias: information comes from study assessed as unclear risk of bias for the majority of domains.

c. Downgraded one level for imprecision: potentially insufficient sample.

Legend

● Low risk of bias.

● Unclear risk of bias.

● High risk of bias.

Newcastle-Ottawa quality assessment for cross-sectional studies

Study	Selection				Comparability	Outcome		Total score
	Representativeness of sample	Sample size	Non-respondents	Ascertainment of exposure		Assessment of outcome	Appropriateness of statistical test	
Possible score	1	1	1	2	2	1	2	Max: 10
Bagonza, Kilbira and Rutebemberwa *(1)*	✹	✹		✹		✹	✹	5

Evidence to decision table

Recommendation 8

*WHO **recommends** providing paid CHWs with a written agreement specifying role and responsibilities, working conditions, remuneration and workers' rights.*

Certainty of the evidence – very low. Strength of the recommendation – strong.

Population: practising CHWs

Intervention: formal contract vs not

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ● Large ○ Varies ○ Don't know	• The systematic review team included two quantitative studies suggesting that CHW performance may be improved as a consequence of formal contracts
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• The systematic review did not find any studies examining any harmful or unintended consequences of formal contracts for CHWs. However, unforeseen legal or administrative implications should be considered as factors that might possibly create resistance to or limit the application of the recommendation
Certainty of evidence	● Very low ○ Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as very low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ○ Probably favours the intervention ● Favours the intervention ○ Varies ○ Don't know	• The GDG was of the view that the benefits of formal contracts, in terms of better labour rights and CHW motivation and performance, outweigh any potential unintended consequences of administrative nature

Resource use and cost-effectiveness	○ Large costs ○ Moderate costs ● Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• No evidence was found on resource requirements or cost-effectiveness • While the costs of paying CHWs are likely to be large (see Annex 6.7), the additional costs of formalizing contractual arrangements were deemed by the GDG to be negligible
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The GDG was of the view that protecting the labour rights of CHWs would improve their motivation, retention and performance, ultimately resulting in positive impacts on health equity
Acceptability and feasibility of intervention	○ No ○ Probably no ○ Probably yes ● Yes (feasibility) ● Varies (acceptability) ○ Don't know	• The stakeholder perception survey found formal contracting of CHWs to be broadly acceptable and feasible • The GDG was of the view that the acceptability of this policy option by policy-makers could vary by country context

Annex 6.8 references

1. Bagonza J, Kilbira SPR, Rutebemberwa E. Performance of community health workers managing malaria, pneumonia and diarrhoea under the community case management programme in central Uganda: a cross sectional study. Malaria Journal. 2014;13(367):1–10.

6.9 *Career ladder.* In the context of CHW programmes, should practising CHWs have a career ladder opportunity or framework versus not?

Overview of quantitative findings

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
Intervention studies							
Ashraf, Bandiera and Lee *(1)*	Cluster RCT	Zambia: recruitment of community health assistants (CHAs) into a role whereby they will receive one year of training, then primarily undertake a role that involves household visits providing advice on women's, children's and environmental health	CHAs recruited from a district where the role was advertised as having career and promotion benefits	CHAs recruited from a district where the role was advertised as having community benefits similar to existing roles	Primary outcome: number of household visits undertaken by CHAs Secondary outcome: use of health service by women and children	Number of visits undertaken by CHAs: those CHAs in the intervention group undertook 29% more household visits than those in the comparison group Women's and children's use of health services: number of women giving birth in a health centre increased by 31% compared to baseline, as did the number of children aged under 5 years visited (increased by 24%), the number of children under 5 years weighed (increased by 23%), and the number of children under 12 months receiving polio vaccination (increased by 20%)	**Results:** Outputs Outcomes **Classification:** Direct CHW-attributable changes in health in individual clients **Measures:** Service delivery Health care-seeking behaviour

Study	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective outcome reporting	Other sources of bias	Overall risk of bias
Ashraf, Bandiera and Lee *(1)*	●	●	●	●	●	●	●	●

Legend

● Low risk of bias
● Unclear risk of bias
● High risk of bias

GRADE quality assessment

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Publication bias	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)
Number of home visits completed by CHAs	Outputs	Direct	Service delivery	307	1	18 months	●	No	No	Yes	No	No [a]	No	No	No	⨁⨁ [b,c]
Number of women giving birth in a health centre	Outcomes	CHW-attributable changes among individual clients	Health care-seeking behaviour	307	1	18 months	●	No	No	Yes	No	No [a]	No	No	No	⨁⨁ [b,c]
Number of children aged under 5 years visiting health post	Outcomes	CHW-attributable changes among individual clients	Health care-seeking behaviour	307	1	18 months	●	No	No	Yes	No	No [a]	No	No	No	⨁⨁ [b,c]
Number of children aged under 5 years weighed at health post	Outcomes	CHW-attributable changes among individual clients	Health care-seeking behaviour	307	1	18 months	●	No	No	Yes	No	No [a]	No	No	No	⨁⨁ [b,c]
Number of children aged under 12 months receiving polio vaccination at health post	Outcomes	CHW-attributable changes among individual clients	Health care-seeking behaviour	307	1	18 months	●	No	No	Yes	No	No [a]	No	No	No	⨁⨁ [b,c]

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Notes on GRADE scores

a. Cannot be assessed as outcome is for one study only.

b. Downgraded one level: information comes from study assessed as unclear risk of bias for majority of domains.

c. Downgraded one level: potentially insufficient sample.

Legend

● Low risk of bias.

● Unclear risk of bias.

● High risk of bias.

Evidence to decision table

Recommendation 9

*WHO **suggests** that a career ladder should be offered to practising CHWs, recognizing that further education and career development are linked to selection criteria, duration and contents of pre-service education, competency-based certification, duration of service and performance review.*

Certainty of the evidence – low. Strength of the recommendation – conditional.

Population: practising CHWs

Intervention: offering a career ladder vs not

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ● Large ○ Varies ○ Don't know	• The systematic review identified only one eligible study with some evidence of effectiveness of offering a career ladder for CHWs. The review of reviews provided some indirect supportive evidence of the importance of career ladders, which was also supported by the GDG based on their own experience
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• The depletion of the pool of practising CHWs is a theoretical undesirable effect that can be offset by scaling up training accordingly. Less known are the potential drawbacks in terms of legal and regulatory challenges to implement career ladder schemes for CHWs
Certainty of evidence	○ Very low ● Low ○ Moderate ○ High ○ No included studies	• The systematic review identified one eligible study; the certainty of the evidence was rated as low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes

Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ● Probably favours the intervention ○ Favours the intervention ○ Varies ○ Don't know	• The GDG was of the view that the potential benefits of a career ladder for CHWs in terms of upward social mobility and improved motivation and retention of practising CHWs outweigh the potential harms and concerns on feasibility
Resource use and cost-effectiveness	○ Large costs ● Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• The moderate costs likely to be required for this policy option are justified in light of its potential advantages • No cost-effectiveness evidence was found on this aspect
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• Recruitment of CHWs from the community (see Annex 6.1) combined with career ladders that create upward social mobility have the potential to improve equity
Acceptability and feasibility of intervention	○ No ○ Probably no ● Probably yes (acceptability) ○ Yes ● Varies (feasibility) ○ Don't know	• The stakeholder perception survey found that offering CHWs a career ladder opportunity is acceptable, but its feasibility might be variable across different contexts

Annex 6.9 references

1. Ashraf N, Bandiera O, Lee SS. Do-gooders and go-getters: career incentives, selection, and performance in public service delivery. Discussion paper. Harvard University; 2014.

6.10 *Target population size.* In the context of CHW programmes, should there be a target population size versus not?

Quantitative findings: overview of included quantitative studies

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
Intervention studies							
Kalyango et al. *(1)*	RCT	Villages in eastern Uganda (n = 65). CHWs, also known as community medicine distributors (n = 125), complemented health services provided by government and NGO health facilities, drug shops and private clinics	Intervention CHWs managed both malaria and pneumonia in children aged under 5 years	Comparison CHWs managed malaria alone in children aged under 5 years	Performance was assessed using knowledge tests, case scenarios of sick children, review of CHW registers and observation of CHWs	Both dual and single management arms (respectively) had similar performance in: • knowledge of malaria (72% vs 70%) • eliciting malaria signs and symptoms (50% in both) • prescription of antimalarials in case scenarios (82% vs 80%) • correct prescription of antimalarials from record reviews (99% vs 100%) Scores differed between dual vs single on: • overall knowledge (72% vs 40%, P < 0.001) • correct doses of medicine from records (100% vs 96%, P < 0.001)	Knowledge, competency, service delivery
Observational studies							
Kuule et al. *(2)*	Cross-sectional study	Catchment of Bwindi Community Hospital in south-western Uganda. Information was collected on community health volunteers (CHVs) to assess their work output via sociodemographic and workplace characteristics (n = 508)	NA	NA	Frequencies and proportions were reported for characteristics. To assess work outputs, study-specific targets were defined	37% of CHVs took care of more than the recommended number of households Overseeing more than the recommended number of households reduced overall performance of CHVs (adjusted OR 0.6; 95% CI, 0.4–0.9, P = 0.02) A “medium” workload (20–30 households) yielded significantly better “household follow-up and reporting” (adjusted OR 0.6; 95% CI, 0.4–0.9) and “malnutrition screening” (adjusted	Service delivery

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
						OR 0.6; 95% CI, 0.4–1.0) when compared to a greater workload (> 30 households), while the effect on "meeting attendance", "immunization coverage" and "deliveries in health unit" were similar between the two groups	
Maji et al. *(3)*	Cross-sectional study	Two districts of West Bengal. Female health workers (FHWs) (n = 26) that offered a wide range of services were sampled	NA	NA	Performance for three indicators (diphtheria-tetanus-pertussis (DTP) booster coverage, antenatal check-up coverage and family planning performance)	CHWs covering smaller populations (less than 6000) showed poorer performance in vaccine coverage CHWs working with larger populations (more than 6000) had higher rates of family planning coverage	Service delivery
Sadler et al. *(4)*	Prospective cohort	Barisal division in southern Bangladesh. CHWs (n = 724) were recruited	All children aged over 6 months identified as suffering from severe acute malnutrition were treated by the CHWs	In comparison group, children identified with severe acute malnutrition were directly send to treatment at health facilities	The feasibility and effectiveness of adding severe acute malnutrition to CHW responsibilities was assessed	A high recovery rate was noted in the intervention group (92%) Mortality was low in the intervention group (0.1%) and not measured in the inpatient group Overall, CHW management of severe acute malnutrition was of high quality, with 58.2% of the sample achieving a perfect score of 100% error-free case identification and management	Service delivery/impact
Suri, Gan and Carpenter *(5)*	Survey	Outer West district of Durban municipality in KwaZulu-Natal, South Africa. CHWs (n = 125) were recruited, whose job responsibilities included disease surveillance, education and counselling	NA	NA	A 30-question written questionnaire assessing CHW perspectives on HIV/AIDS, TB, and potential CHW programme improvement was administered	Approximately 101 households (mean 101.04; 95% CI, 91.62–110.46) were allocated to each CHW For households with TB, 84% of CHWs supported that there should be a daily visit to ensure compliance with treatment (DOTS-consistent notion), while CHWs reported visiting a mean of five households/day (95% CI, 4.93–5.77 households/day)	Service delivery

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
Definitions **Odds ratio (OR).** A measure of effect that is used to approximate relative risk (i.e., the likelihood that one group will experience the outcome given a certain exposure versus the likelihood that another group will experience the outcome given they were not exposed). When the OR is greater than 1.0, the risk is greater. When the OR is between 0 and 1, the risk is lower. When the risk is 1.0, there is no difference between groups. The further the OR is above or below 1.0, the larger the effect. **Correlation.** A measure of association between two different constructs. **Significance or statistical significance.** The probability that a finding was observed by chance alone. Traditionally, a finding is said to be "significant" when this probability is less than 0.05 (i.e., $P < 0.05$). **Confidence interval (CI).** The estimated interval between which the measure of effect (e.g. the OR) would probably be observed if the study were conducted again on a similar sample of subjects. **Adjusted (e.g. adjusted OR).** When an explanatory or causal factor's raw association with an outcome is statistically adjusted to take account of other potential explanatory factors.							

GRADE quality assessment

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Other considerations	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)
Cost saving																
Cost per disability-adjusted life-year (DALY) averted in task shifting (US$) (non-RCT)	Outcome	CHW-attributable changes among individual clients	Cost saving	724	1: Das et al. *(6)*		🟡	[a]	No	No	[a]	No	No	No	No	⨁ [b]
Service delivery																
Overall performance of CHWs based on number of households (non-RCT)	Outcome	CHW-attributable changes among individual clients	Service delivery	508	1: Das et al. *(6)*		🟢	[a]	No	No	[a]	No	No	No	No	⨁⨁
CHW performance for vaccinations, antenatal care and contraceptive use based on population size (non-RCT)	Outcome	CHW-attributable changes among individual clients	Service delivery	42	1: Maji et al. *(3)*		🟢	[a]	No	Yes	[a]	No	No	No	No	⨁ [c]
Mean number of household visits/day by CHW	Outcome	CHW-attributable changes among individual clients	Service delivery	120	1: Suri, Gan and Carpenter *(5)*		🟢	[a]	No	Yes	[a]	No	No	No	No	⨁ [c]
Impact																
Morbidity/mortality (RCT)	Outcome	Direct	Impact	724	1: Chang et al. *(7)*		🔴	[a]	No	Yes	[a]	No	No	No	No	⨁⨁ [b,d]

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Other considerations	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)

Notes

⊕ indicates that the overall quality of evidence is very low.
⊕⊕ indicates that the overall quality of evidence is low.
⊕⊕⊕ indicates that the overall quality of evidence is moderate.
⊕⊕⊕⊕ indicates that the overall quality of evidence is high.

Notes on GRADE scores

a. Only one study for the outcome, thus cannot be assessed.
b. Downgraded one level for risk of bias: information comes from studies assessed as unclear or high risk of bias for the majority of domains.
c. Downgraded one level: potentially insufficient sample.
d. Downgraded one level for imprecision: wide confidence interval.

Legend

- Low risk of bias.
- Unclear risk of bias.
- High risk of bias.

Newcastle-Ottawa quality assessment for cross-sectional studies

	Selection				Comparability	Outcome		
Study	**Representativeness of sample**	**Sample size**	**Non-respondents**	**Ascertainment of exposure**		**Assessment of outcome**	**Appropriateness of statistical test**	**Total score**
Possible score	1	1	1	2	2	1	2	Max: 10
Kuule et al. *(2)*	★	★		★	★★	★	★	7
Suri, Gan and Carpenter *(5)*	★	★		★	★★	★	★	7
Sadler et al. *(4)*	★	★		★	★	★	★	6
Maji et al. *(3)*	★	★		★	★★	★	★	7

Evidence to decision table

Recommendation 10

*WHO **suggests** using the following criteria in determining a target population size in the context of CHW programmes.*

Criteria to be adopted in most settings:

- expected workload based on epidemiology and anticipated demand for services;
- frequency of contact required;
- nature and time requirements of the services provided;
- expected weekly time commitment of CHWs (factoring in time away from service provision for training, administrative duties, and other requirements);
- local geography (including proximity of households, distance to clinic and population density).

Criteria that might be of relevance in some settings:

- weather and climate;
- transport availability and cost;
- health worker safety;
- mobility of population;
- available human and financial resources.

Certainty of the evidence – very low. Strength of the recommendation – conditional.

Population: practising CHWs

Intervention: use of target population size vs not

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• The systematic review team findings suggest that CHW performance is influenced by the population size or workload that is assigned to them. However, evidence on the effectiveness of an optimal population size for CHWs is ambiguous, in that some studies suggest that an increased population size or workload compromises CHW performance, while other studies point to the opposite
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ○ Varies ● Don't know	• The systematic review did not find any studies examining any harmful or unintended consequences of having a target population size for CHW programmes
Certainty of evidence	● Very low ● Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as very low

Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ● Probably favours the intervention ○ Favours the intervention ○ Varies ○ Don't know	• The GDG was of the view that there may be more advantages than harms in determining at the national level and on the basis of the suggested criteria a target population size for CHW programmes
Resource use and cost-effectiveness	○ Large costs ● Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• The moderate costs likely to be required for collection of data to inform reliable planning of CHWs can be justified by the improved design and performance of the programmes • The systematic review team identified one study suggesting that increasing the caseload is cost-effective
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The GDG was of the view that defining an appropriate target population size may improve health equity
Acceptability and feasibility of intervention	○ No ○ Probably no ○ Probably yes ● Yes ○ Varies ○ Don't know	• The systematic review team did not identify studies assessing acceptability and feasibility of the policy option under consideration • The GDG was of the view that adopting a target population size will be feasible and acceptable in most contexts

Annex 6.10 references

1. Kalyango JN, Rutebemberwa E, Alfven T, Ssali S, Peterson S, Karamagi C. Performance of community health workers under integrated community case management of childhood illnesses in eastern Uganda. Malaria Journal. 2012;11(1):282.

2. Kuule Y, Dobson AE, Woldeyohannes D, Zolfo M, Najjemba R, Edwin BMR et al. Community health volunteers in primary healthcare in rural Uganda: factors influencing performance. Frontiers in Public Health. 2017;5.

3. Maji D, Hutin Y, Ramakrishnan R, Hossain S, De S. Strategies to improve the performance of female health workers in West Bengal: a cross-sectional survey. National Medical Journal of India. 2010;23(3):137–42.

4. Sadler K, Puett C, Mothabbir G, Myatt M. Community case management of severe acute malnutrition in southern Bangladesh. Feinstein International Center; 2011.

5. Suri A, Gan K, Carpenter S. Voices from the field: perspectives from community health workers on health care delivery in rural KwaZulu-Natal, South Africa. Journal of Infectious Diseases. 2007;196(Suppl. 3):S505–11.

6. Das A, Friedman J, Kandpal E, Ramana GN, Gupta RK, Pradhan MM et al. Strengthening malaria service delivery through supportive supervision and community mobilization in an endemic Indian setting: an evaluation of nested delivery models. Malaria Journal. 2014;13:482.

7. Chang LW, Kagaayi J, Arem H, Nakigozi G, Ssempijja V, Serwadda D 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):1776–84.

6.11 *Data collection and use.* In the context of CHW programmes, should practising CHWs collect, collate, and use health data versus not?

Quantitative findings: overview of included quantitative studies

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
Intervention studies							
Chang et al. *(1)*	Cluster RCT/mixed methods	Uganda: HIV+ peer health workers (PHWs) providing care to patients (n = 970) living with HIV on antiretroviral therapy (ART)	PHWs received training + mHealth support intervention to record patient outcomes (four clusters; n = 13 PHWs; n = 446 patients) PHWs sent an SMS to their supervisors reporting patient "pill count" (ART adherence) and other clinical data	PHWs received training (six clusters; n = 16 PHWs; n = 524 patients)	Primary outcome: cumulative risk of virological failure[2] among patients on ART Secondary outcomes: patient treatment adherence; virological failure at 24 and 48 weeks, loss to follow-up and mortality	Virological failure occurred in 19.4% of intervention group and 16.4% of control group: RR 1.17; 95% CI, 0.84–1.64, P = 0.34 Mortality occurred in 8.3% of intervention group and 10.1% of control group: RR 0.82; 95% CI, 0.55–1.22, P = 0.33	Results: impact Classification: CHW-attributable changes in health at the population level Measures: mortality
DeRenzi *(2)*	RCT	Dar es Salaam, United Republic of Tanzania: CHWs providing monthly visits to patients with chronic conditions (HIV, diabetes, TB, etc.) over 12 weeks (n = 74)	CHWs collected referral and follow-up data through an mHealth application – CommCare. These were automatically stored on a server and triggered SMS reminders to CHWs. A message was sent if they were overdue with a follow-up. If a CHW was more than three days overdue, a message was also sent to a supervisor (n = 34)	No SMS reminders were sent to CHWs (n = 40)	Number of days CHWs were overdue to visit patients	Intervention group: 86% decrease in average number of days that clients are overdue (9.7 days to 1.4 days) from baseline to endline Control group: 13.4% increase in average number of days that clients are overdue (8.2 days to 9.3 days) from baseline to endline The SMS + supervisor intervention significantly reduced the number of days that clients were overdue (U = 271.00, P < 0.001, r =.500)	Results: outputs Classification: direct Measures: absenteeism
Vallières et al. *(3)*	RCT	Bonthe district, Sierra Leone: CHWs providing	Mobile Technology for Community Health	Third group received maternal and child health	Self-reported surveys measuring:	No statistically significant differences in scores were	Results: outputs

[2] A type of HIV treatment failure, virological failure occurs when ART fails to suppress and sustain a person's viral load to less than 200 copies/ml. Factors that can contribute to virological failure include drug resistance, drug toxicity, and poor treatment adherence.

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
		maternal and child health services (n = 313)	(MoTeCH) was tested using a three-armed RCT: First group received maternal and child health training + mobile phone + closed user group (n = 110) Second group received maternal and child health training + mobile phone with MoTeCH + closed user group (n = 94) MoTeCH assists CHWs to register pregnant women and children for services, collect household data, make referrals and receive reminders when appointments are overdue	training + no mobile phone (n = 109)	• supervision (perceived supportive supervision scale) • motivation (volunteer functions inventory) • work engagement (Utrecht work engagement scale) • job satisfaction (Minnesota satisfaction questionnaire) Surveys were completed at three time points over 18 months	recorded over the three time periods Perceived supervision: (F[2,286] = 0.294, P = .745) Motivation: (F[2,289] = 0.700, P = .497) Work engagement: (F[2,286] = 0.041, P = .960) Job satisfaction: (F[2,285] = 1.740, P = .177)	Classification: indirect Measures: motivation; satisfaction
Observational studies							
McNabb et al. *(4)*	Cohort: pre-post	Nigeria: community health extension workers working in 20 primary health care centres; n = 152 CHWs + 20 supervisors	Use of a mobile case management and decision support application (CommCare) guides community health extension workers through antenatal care procedures and collects patient data	Not applicable	Quality of antenatal care services measured by a 25-indicator quality score collected at baseline and endline Clients surveyed: n = 267	Quality score: 13.3 (baseline), 17.2 (endline) ($P < 0.0001$) Client satisfaction: 75.4% (satisfied at baseline) to 83.3% (satisfied at endline) ($P < 0.025$)	Results: outputs Classification: direct Measures: service delivery
Oum, Chandramohan and Cairncross *(5)*	Cohort: pre-post	Cambodia: Pir Thnu, Preah Rumkei, Trang, Ta Saen Boeng Reang communes. Village health volunteers (VHVs) using a community-based surveillance system	VHVs use a community-based surveillance system to report data on suspected outbreaks, infectious diseases and vital events	Not applicable	Criteria: • outbreak detection • monitoring of suspected outbreaks • vital events • accuracy of reporting	Infant mortality: 80.0 per 1000 live births (baseline) to 72.9 per 1000 live births (endline) Under-5 mortality: 107.5 per 1000 live births (baseline) to 89.0 per 1000 live births (endline)	Results: impact Classification: CHW-attributable changes in health at the population level Measures: morbidity

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
Shieshia et al. *(6)*	Cohort: pre-post non-equivalent control group	Malawi, 10 out of 28 districts selected with a functioning integrated community case management (iCCM) system. Health surveillance assistants and health facility staff delivering iCCM	The mHealth intervention being tested (cStock) is a tool for community-level reporting on available health stocks and resupplying 19 items used by health surveillance assistants. The intervention was tested under three conditions: 1. In three districts, cStock was used by health surveillance assistants and health facility staff alongside enhanced management (EM) 2. In three districts, cStock was used by health surveillance assistants and health facility staff with efficient product transport (EPT)	3. No intervention (four districts)	Criteria: • feasibility (mean stock reporting rate) • acceptability (mean reporting completeness) • effectiveness (lead time for drug resupply; stock-out rates)	The study only reports results from six cStock sites and thus only compares EM and EPT sites: • mean stock reporting rate was 94% in EM and 79% in EPT (t = 6.9766, $P < 0.001$) • mean reporting completeness was 85% in EM and 65% in EPT (t = 9.8953, $P < 0.001$) • lead time for drug resupply was 12.8 days in EM group and 26.4 days in EPT group (t = 7.75, $P < 0.001$) • mean stock-out rate was significantly lower in EM group (5–7%) compared to EPT group (10–21%) ($P < 0.001$)	Results: outputs Classification: CHW-attributable changes in the health system Measures: change in health system functioning
Bagonza, Kilbira and Rutebemberwa *(7)*	Cross-sectional	Wakiso district, Uganda: community health workers providing integrated care and case management ($n = 336$)	CHWs providing care under the iCCM programme were surveyed to determine if there were factors that assessed the quality of care they provided. The "regular submission of monthly reports" is of interest to this review	Not applicable	Data were collected through interviews with CHWs and reviewing health record collecting through the iCCM programme. Performance was assessed through composite scores. To assess data collection, "regular submission of monthly reports" was examined	CHWs categorized as "good performers" were 5.75 more likely to submit data reports when compared with CHWs categorized as "low performers" Performance was assessed using composite scores: "regular submission of monthly reports" (unadjusted OR 5.75; 1.70–19.34)	Results: outputs Classification: direct Measures: service delivery
Kuhn and Zwarenstein *(8)*	Cross-sectional	South Africa: Thornhill, a rural village of 12 000 people with a village health worker	VHWs were introduced to a paper-based record system that allowed them to record and track demographic and health information about newborn children in the village. The records	Not applicable	Criteria: • presence of VHW card • use of breastfeeding • immunization rates	Breastfeeding at 11 months: 87.3% (VHW card group) vs 78% (no VHW card) ($P = 0.0045$) Polio vaccination: 66.7% (VHW card group) vs 50.3% (no VHW card) ($P = 0.0196$)	Results: outcomes Classification: CHW-attributable changes in the community Measures:

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
		programme; *n* for survey not stated	included immunization, number of visits and information covered in CHW visits			Measles vaccination: 33.2% (VHW card group) vs 60.5% (no VHW card) ($P = 0.0001$)	change in community health
Ngabo et al. *(9)*	Cross-sectional	Musanze district, Rwanda: CHWs tasked with maternal and child health ($n = 432$)	Rapid SMS-MCH – a tool for registering and monitoring pregnancies, reporting threats to maternal health and providing guidelines to CHWs – was used by 432 CHWs in the district	Not applicable	Criteria: • count of use: number of reports submitted • % of pregnancies registered • reporting compliance	Count of use: 35 734 reports submitted Approximately 81% of pregnancies registered Reporting compliance for all registered pregnancies was 100%	Results: outputs Classification: direct Measures: service delivery
Umar, Olumide and Bawa *(10)*	Cross-sectional	Akinyele LGA, Oyo state, Nigeria: voluntary village health workers (VHWs) were surveyed to obtain their views on the collection of data to inform the Primary Health Care Management Information System ($n = 102$)	Trained VHWs and traditional birth attendants in study area were eligible if they had undertaken training in their role, been issued with kits and been identified as active by their supervisor in the last 12 months ($n = 102$). These VHWs and traditional birth attendants were asked about their data collection and record-keeping practices and attitudes towards data collection	Not applicable	Semi-structured questionnaire and observation checklist administered by research assistants. Response rate not included	VHWs who: • kept health records: 96.1% • forwarded health records to district: 95.9% • perceived record keeping as "easy": 93.1% Several other data points included in the study	Results: outputs Classification: indirect Measures: knowledge

Summary of findings

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
Virological failure amongst patients receiving antiretroviral treatment for HIV/AIDS			1.17; 0.84–1.64	29 (1)	⨁ [a]
Immunization rates for polio and measles at 11 months following intervention commencement	50.3%	66.7%	Not possible to determine	228 (1)	⨁ [b]
Breastfeeding rates at 11 months following intervention commencement	78%	87.3%	Not possible to determine	228 (1)	⨁ [b]
Composite score measuring performance of CHW	NA	NA	5.75; 1.70–19.34	336 (1)	⨁ [c]
Infant mortality in villages where surveillance system is used	80 per 1000 live births	72.9 per 100 live births	Not possible to determine	52 (1)	⨁ [b]
Under-5 mortality in villages where surveillance system is used	107.5 per 1000 live births	89.0 per 1000 live births	Not possible to determine	52 (1)	⨁ [b]
Quality of antenatal services measured by a quality score	13.33	17.15	Not possible to determine	267 (1)	⨁ [c]

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Summary findings

a. There is limited evidence that CHW using data can lead to CHW-attributable changes at the population level.

b. There is limited evidence that CHW using data can lead to CHW-attributable changes among individual clients.

c. There is limited evidence that CHW using data can lead to changes in service delivery.

GRADE quality assessment

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)
Virological failure amongst patients receiving antiretroviral treatment for HIV/AIDS	Impact	CHW-attributable changes at the population level	Mortality	29	1: Chang et al. *(1)*	2 years	🟡	No serious inconsistency[a]	Serious indirectness[1]	Serious imprecision[1]	Undetected[b]	No	No	No	⊕
Immunization rates for polio and measles at 11 months following intervention commencement	Outcomes	CHW-attributable changes among individual clients	Health	228	1: Kuhn and Zwarenstein *(8)*	NA	🔴	No serious inconsistency	No serious indirectness	Serious imprecision[2]	Undetected	No	No	No	⊕
Breastfeeding rates at 11 months following intervention commencement	Outcomes	CHW-attributable changes among individual clients	Health	228	1: Kuhn and Zwarenstein *(8)*	NA	🔴	No serious inconsistency	No serious indirectness	Serious imprecision[2]	Undetected	No	No	No	⊕
Composite score measuring performance of CHW	Outputs	Direct	Service delivery	336	1: Bagonza, Kilbira and Rutebemberwa *(7)*	NA	🔴	No serious inconsistency	No serious indirectness	Serious imprecision[3]	Undetected	No	No	No	⊕
Infant mortality in villages where surveillance system is used	Impact	CHW-attributable changes in health at the population level	Morbidity	52	1: Oum, Chandramohan and Cairncross *(5)*	2 years	🔴	No serious inconsistency	Serious indirectness[4]	Serious imprecision[4]	Undetected	No	No	No	⊕
Under-5 mortality in villages where surveillance system is used	Impact	CHW-attributable changes in health at the population level	Morbidity	52	1: Oum, Chandramohan and Cairncross *(5)*	2 years	🔴	No serious inconsistency	Serious indirectness[4]	Serious imprecision[4]	Undetected	No	No	No	⊕
Quality of antenatal services measured by a quality score	Outputs	Direct	Service delivery	267	1: McNabb et al. *(4)*	1 year	🔴	No serious inconsistency	No serious indirectness	Serious imprecision[5]	Undetected	No	No	No	⊕

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Legend

- Low risk of bias.
- Unclear risk of bias.
- High risk of bias.

Notes on GRADE scores and sources

a. When only one study contributes evidence to an outcome, no serious inconsistency is assumed.

b. Undetected is recorded for publication bias, as it cannot be assessed for one study.

[1] Chang et al. *(1)*: *serious indirectness*, as both the intervention and comparison group include the collection of health data. The intervention tests the effect of the method of data collection (paper versus electronic) on patient outcomes. *Serious imprecision*, due to the small number of clusters, participants, and events for most study outcomes. While RCTs start out as high, this study receives a final GRADE assessment of *very low* after being rated down for unclear risk of bias, serious indirectness, and serious imprecision.

[2] Kuhn and Zwarenstein *(8)*: *serious imprecision*, due to an inadequate number of events (< 300) for the outcome. As this is a non-RCT, which starts out as low, this study receives a final GRADE assessment of *very low* after being rated down for high risk of bias and serious imprecision.

[3] Bagonza, Kilbira and Rutebemberwa *(7)*: *serious imprecision*, as the confidence interval is large and the sample size by event (e.g., the proportion who have regular submission of monthly reports by performance level) is not reported. However, it is assumed that these are both under 300 as the sample size in totality is 336. As this is a non-RCT, which starts out as low, this study receives a final GRADE assessment of *very low* after being rated down for high risk of bias and serious imprecision.

[4] Oum, Chandramohan and Cairncross *(5)*: *serious indirectness*, as there are no study outcomes that assess the impact of the surveillance system on CHW performance. *Serious imprecision*, as there is an inadequate number of events per outcome. As this is a non-RCT, which starts out as low, this study receives a final GRADE assessment of *very low* after being rated down for high risk of bias, serious indirectness, and serious imprecision.

[5] McNabb et al. *(4)*: *serious imprecision*, based on the rule of thumb of 400 participants for adequate precision with continuous outcomes. As this is a non-RCT, which starts out as low, this study receives a final GRADE assessment of *very low* after being rated down for high risk of bias and serious imprecision.

Newcastle-Ottawa quality assessment for cohort and cross-sectional studies

	Selection				Comparability	Outcome			
Study	**Representativeness of exposed sample**	**Selection of the exposed sample**	**Ascertainment of exposure**	**Demonstration that outcome of interest was not present at start of study**	**Comparability of cohorts on basis of the design or analysis**	**Assessment of outcome**	**Was follow-up long enough for outcomes to occur**	**Adequacy of follow-up of cohorts**	**Total score**
Possible score	1	1	1	1	2	1	1	1	Max: 9
McNabb et al. *(4)*						✹			1
Oum, Chandramohan and Cairncross *(5)*		✹	✹		✹		✹		4
Shieshia et al. *(6)*	✹	✹	✹	✹	✹	✹			6

	Selection				Comparability	Outcome		
Study	**Representativeness of sample**	**Sample size**	**Non-respondents**	**Ascertainment of exposure**		**Assessment of outcome**	**Appropriateness of statistical test**	**Total score**
Possible score	1	1	1	2	2	1	2	Max: 10
Bagonza, Kilbira and Rutebemberwa *(7)*	✹	✹		✹		✹	✹	5
Kuhn and Zwarenstein *(8)*	✹	✹		✹	✹			4
Ngabo et al. *(9)*	✹							1
Umar, Olumide and Bawa *(10)*				✹		✹		2

Qualitative findings

Objective	To identify, appraise and synthesize qualitative research evidence on the potential benefits of CHWs using data collection as part of their regular work routines		
Perspective	CHW experiences of the benefits of using data		
Included programmes	Studies that assess the regular use of data by CHWs		
Review finding	**Overall CERQual assessment of confidence**	**Explanation of CERQual assessment**	**Studies contributing to the review**
Service delivery: Data collection leads to a greater understanding of client's health and needs	Low confidence	This finding was graded as low confidence due to minor methodological limitations, moderate concerns about relevance, and substantial concerns regarding adequacy	Zanchetta et al. *(11)*
Productivity: Data collection using mHealth leads to increased efficiency in routine work	Moderate confidence	This finding was graded as moderate confidence due to minor methodological limitations, moderate concerns about relevance, high confidence in coherence and moderate confidence in adequacy	Madon et al. *(12)*, Braun et al. *(13)*
Change in health system functioning: Data collection using mHealth leads to more accurate communication between CHWs and supervisors	Moderate confidence	This finding was graded as moderate confidence due to minor methodological limitations, moderate concerns about relevance and moderate confidence in adequacy	Madon et al. *(12)*, Braun et al. *(13)*
Self-efficacy/esteem: Data collection using mHealth leads to increased motivation and self-esteem	Moderate confidence	This finding was graded as moderate confidence due to minor methodological limitations, moderate concerns about relevance and moderate confidence in adequacy	Madon et al. *(12)*
Credibility: Data collection using mHealth leads to increased credibility of CHWs amongst community population	Moderate confidence	This finding was graded as moderate confidence due to minor methodological limitations, moderate concerns about relevance, high confidence in coherence and moderate confidence in adequacy	Madon et al. *(12)*, Braun et al. *(13)*
Attrition: Collection of data by CHWs could lead to higher retention rates due to increased level of engagement	Low confidence	This finding was graded as low confidence due to significant concerns about methodology	Strachan et al. *(14)*

Evidence to decision table

Recommendation 11

*WHO **suggests** that practising CHWs document the services they are providing and that they collect, collate and use health data on routine activities, including through relevant mobile health solutions. Enablers for success include minimizing the reporting burden and harmonizing data requirements; ensuring data confidentiality and security; equipping CHWs with the required competencies through training; and providing them with feedback on performance based on data collected.*

Certainty of the evidence – very low. Strength of the recommendation – conditional.

Population: practising CHWs

Intervention: collection, collation and use of data by CHWs vs not

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ● Large ○ Varies ○ Don't know	• Several studies were identified suggesting a positive potential of CHW involvement in data collection, collation and use
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• The additional burden posed on CHWs by data collection requirements was flagged
Certainty of evidence	● Very low ○ Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as very low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ● Probably favours the intervention ○ Favours the intervention ○ Varies ○ Don't know	• The GDG was of the view that the benefits of CHW involvement in data collection and use outweigh potential harms, if the enabling factors for success mentioned in the recommendation are put in place

Resource use and cost-effectiveness	● Large costs ○ Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• The costs are likely to be significant but justified in light of the potential benefits. The systematic review of reviews identified one review pointing to potential savings through mHealth applications compared to traditional methods of data collection and transmission
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The GDG was of the view that embedding equity considerations in the design of data collection by CHWs might be instrumental to improving health equity
Acceptability and feasibility of intervention	○ No ○ Probably no ● Probably yes ○ Yes ○ Varies ○ Don't know	• The systematic review team identified one study indicating ambiguity, since CHWs described data collection as both acceptable and burdensome. Another study identified data collection processes as having the potential to increase the social status of CHWs, since data were collected through technologies provided through government • The systematic review of reviews found that district health teams may be sceptical about the value and quality of data collected by CHWs • The stakeholder perception survey found data collection and use by CHWs to be both acceptable and feasible

Annex 6.11 references

1. Chang LW, Njie-Carr V, Kalenge S, Kelly JF, Bollinger RC, Alamo-Talisuna S. Perceptions and acceptability of mHealth interventions for improving patient care at a community-based HIV/AIDS clinic in Uganda: a mixed methods study. AIDS Care. 2013;25(7):874–80.
2. DeRenzi B. Technology for workforce performance improvement of community health programs [dissertation]. Seattle: University of Washington; 2011.
3. Vallières F, McAuliffe E, van Bavel B, Wall P, Trye A. There's no app for that: assessing the impact of mHealth on the supervision, motivation, engagement, and satisfaction of community health workers in Sierra Leone. Annals of Global Health. 2016;82(5):936–49.
4. McNabb M, Chukwu E, Ojo O, Shekhar N, Gill CJ, Salami H et al. Assessment of the quality of antenatal care services provided by health workers using a mobile phone decision support application in northern Nigeria: a pre/post-intervention study. PLoS ONE. 2015;10(5):1–11.
5. Oum S, Chandramohan D, Cairncross S. Community-based surveillance: a pilot study from rural Cambodia. Tropical Medicine and International Health. 2005;10(7):689–97.
6. Shieshia M, Noel M, Andersson S, Felling B, Alva S, Agarwal S et al. Strengthening community health supply chain performance through an integrated approach: using mHealth technology and multilevel teams in Malawi. Journal of Global Health. 2014;4:2.

7. Bagonza J, Kilbira SPR, Rutebemberwa E. Performance of community health workers managing malaria, pneumonia and diarrhoea under the community case management programme in central Uganda: a cross sectional study. Malaria Journal. 2014;13(367):1–10.

8. Kuhn L, Zwarenstein M. Evaluation of a village health worker programme: the use of village health worker retained records. International Journal of Epidemiology. 1990;19:3.

9. Ngabo F, Nguimfack J, Nwaigwe F, Mugeni C, Muhoza D, Wilson D et al. Designing and implementing an innovative SMS-based alert system (RapidSMS-MCH) to monitor pregnancy and reduce maternal and child deaths in Rwanda. Pan Africa Medical Journal. 2012;13:31.

10. Umar US, Olumide EAA, Bawa SB. Voluntary health workers' knowledge, attitude and practices regarding record keeping in Akinyele LGA of Oyo state, Nigeria. Nigerian Postgraduate Medical Journal. 2002;9(1):17–22.

11. Zanchetta MS, Pinto RM, Galhego-Garcia W, da Cunha Z, Cordeiro HA, Fagundes-Filho FE et al. Brazilian community health agents and qualitative primary healthcare information. Primary Health Care Research and Development. 2016;16:235–45.

12. Madon S, Amaguru JO, Malecela MN, Michael E. Can mobile phones help control neglected tropical diseases? Experiences from Tanzania. Social Science and Medicine. 2014;102:103–10.

13. Braun R, Lasway C, Agarwal S, L'Engle K, Layer E, Silas L et al. An evaluation of a family planning mobile job aid for community health workers in Tanzania. Contraception. 2016;94:27–33.

14. Strachan D, Källander K, ten Asbroek A, Kirkwood B, Meek S, Benton L et al. Interventions to improve motivation and retention of community health workers delivering integrated community case management (iCCM): stakeholder perceptions and priorities. American Journal of Tropical Medicine and Hygiene. 2012;87(Suppl. 5):111–9.

6.12 *Types of CHWs.* In the context of CHW programmes, should practising CHWs work in a multi-cadre team versus in a single-cadre CHW system?

Evidence to decision table

Recommendation 12 *WHO **suggests** adopting service delivery models comprising CHWs with general tasks as part of integrated primary health care teams. CHWs with more selective and specific tasks can play a complementary role when required on the basis of population health needs, cultural context and workforce configuration.* Certainty of the evidence – very low. Strength of the recommendation – conditional.		
Population: practising CHWs Intervention: generalist polyvalent CHWs vs specialized CHWs		
Factors	**Decision**	**Explanations/comments**
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• The systematic review did not identify any study eligible for inclusion
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• The systematic review did not identify any study eligible for inclusion
Certainty of evidence	○ Very low ○ Low ○ Moderate ○ High ● No included studies	• The systematic review did not identify any study eligible for inclusion
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ● Probably favours the intervention ○ Favours the intervention	• The GDG was of the view that a model based on polyvalent CHWs, complemented if needed by additional CHWs focusing on more specialized tasks, represents in most settings the most appropriate approach to maximize benefits and contain possible harms

	○ Varies ○ Don't know	
Resource use and cost-effectiveness	● Large costs ○ Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ● No included studies	• The systematic review did not identify any study eligible for inclusion
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The GDG was of the view that polyvalent CHWs, by responding more holistically to population needs, may be a more suitable approach to improve health equity. Conversely, a model based on specialized CHWs might lead to patients affected by certain conditions having poorer access to treatment than those affected by other conditions
Acceptability and feasibility of intervention	○ No ○ Probably no ● Probably yes ○ Yes ○ Varies ○ Don't know	• The GDG was of the view that a CHW model based on polyvalent CHWs would be acceptable and feasible in most contexts

6.13 *Community engagement.* In the context of CHW programmes, are community engagement strategies effective in improving CHW programme performance and utilization?

Summary of quantitative findings

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants	Quality of evidence (GRADE)
	Control risk	Intervention risk			
Neonatal mortality (rate per 1000 live births)	49.1	43.0	0.85; 0.76–0.96 [a]	23 033	⨁⨁
Stillbirth (rate per 1000 live births)	48.7	39.1	0.79; 0.68–0.92 [a]	23 033	⨁⨁
Maternal mortality rate (women's group vs no women's group in peer counselling)			0.26; 0.1–0.7 [b]	185 888	⨁⨁
Perinatal mortality rate (women's group vs no women's group, no peer counselling)			0.67; 0.50–0.88 [b]	185 888	⨁⨁
Neonatal mortality rate (women's group vs no women's group, no peer counselling)			0.59; 0.40–0.86 [b]	185 888	⨁⨁
Infant mortality rate (women's group vs no women's group, no peer counselling)			0.72; 0.56–0.94 [b]	185 888	⨁⨁
Exclusive breastfeeding (volunteer peer counselling vs no peer counselling in women's group areas)			5.13; 2.55–10.33 [b]	185 888	⨁⨁

Notes

⨁ indicates that the overall quality of evidence is very low.
⨁⨁ indicates that the overall quality of evidence is low.
⨁⨁⨁ indicates that the overall quality of evidence is moderate.
⨁⨁⨁⨁ indicates that the overall quality of evidence is high.

Notes on relative effect

a. Mortality risk ratio.
b. Adjusted odds ratio.

GRADE quality assessment

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)
Neonatal mortality	Impact	CHW-attributable changes in health at the population level	Mortality	23 033	1: Bhutta et al. *(1)*		●	No serious inconsistency[a]	No serious indirectness	No serious imprecision	Undetected[b]	No	No	Yes[1]	⊕⊕⊕
Stillbirth	Impact	CHW-attributable changes in health at the population level	Mortality	23 033	1: Bhutta et al. *(1)*		●	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	No	No	Yes[1]	⊕⊕⊕
Maternal mortality rate (women's group vs no women's group in peer counselling)	Impact	CHW-attributable changes in health at the population level	Mortality	185 888	1: Lewycka et al. *(2)*	1 year	●	No serious inconsistency	Serious indirectness[2]	No serious imprecision	Undetected	Yes[2]	No	Yes[2]	⊕⊕⊕
Perinatal mortality rate (women's group vs no women's group, no peer counselling)	Impact	CHW-attributable changes in health at the population level	Mortality	185 888	1: Lewycka et al. *(2)*	1 year	●	No serious inconsistency	Serious indirectness[2]	No serious imprecision	Undetected	No	No	Yes[2]	⊕⊕
Neonatal mortality rate (women's group vs no women's group, no peer counselling)	Impact	CHW-attributable changes in health at the population level	Mortality	185 888	1: Lewycka et al. *(2)*	1 year	●	No serious inconsistency	Serious indirectness[2]	No serious imprecision	Undetected	No	No	Yes[2]	⊕⊕
Infant mortality rate (women's group vs no women's group, no peer counselling)	Impact	CHW-attributable changes in health at the population level	Mortality	185 888	1: Lewycka et al. *(2)*	1 year	●	No serious inconsistency	Serious indirectness[2]	No serious imprecision	Undetected	No	No	Yes[2]	⊕⊕
Exclusive breastfeeding (volunteer peer counselling vs no peer counselling in women's group areas)	Impact	CHW-attributable changes among individual clients	Health-promoting behaviour in the home	185 888	1: Lewycka et al. *(2)*	1 year	●	No serious inconsistency	Serious indirectness[2]	No serious imprecision	Undetected	Yes[2]	No	Yes[2]	⊕⊕⊕

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)

Notes

⊕ indicates that the overall quality of evidence is very low.

⊕⊕ indicates that the overall quality of evidence is low.

⊕⊕⊕ indicates that the overall quality of evidence is moderate.

⊕⊕⊕⊕ indicates that the overall quality of evidence is high.

Legend

- Low risk of bias.
- Unclear risk of bias.
- High risk of bias.

Notes on GRADE scores and sources

a. When only one study contributes evidence to an outcome, no serious inconsistency is assumed.

b. Undetected is recorded for publication bias, as it cannot be assessed for one study.

[1] Bhutta et al. *(1)*: confounding is plausibly controlled for through randomization and is discussed in the limitations. While RCTs start out as high, this study receives a final GRADE assessment of *moderate* after being rated down for unclear risk of bias.

[2] Lewycka et al. *(2)*: *serious indirectness*, as some of the outcomes compare the women's group with the peer counselling group, both of which have some elements of community engagement. The magnitude of effect warrants grading up for two outcomes: maternal mortality rate (RR 0.26; 0.1–0.7) and exclusive breastfeeding (RR 5.13; 2.55–10.33). Confounding is plausibly controlled for through randomization. While RCTs start out as high, this study receives a final GRADE assessment of *low* for three of five outcomes (perinatal mortality, neonatal mortality, and infant mortality rate), after being rated down for unclear risk of bias and serious indirectness. This study receives a final GRADE assessment of *moderate* for two of five outcomes (maternal mortality rate, exclusive breastfeeding).

Risk of bias: modified Newcastle-Ottawa quality assessment scale for cross-sectional studies

Study	Selection				Comparability	Outcome		Total score
	Representativeness of sample	Sample size	Non-respondents	Ascertainment of exposure		Assessment of outcome	Appropriateness of statistical test	
Possible score	1	1	1	2	2	1	2	Max: 10
Abbey et al. *(3)*		●		●	●	●	●	5
Adams, Nababan and Hanifi *(4)*	●	●	●	●●	●	●	●	8
Adejumo et al. *(5)*					●	●	●	3
Ahluwalia et al. *(6)*				●	●	●		3
Andersen et al. *(7)*	●	●	●	●	●●		●	7
Broadhead et al. *(8)*	●			●	●	●		4
Capell *(9)*				●		●		2
Edward et al. *(10)*	●	●	●	●		●	●	6
Elmardi et al. *(11)*	●			●		●		3
Gopalan, Mohanty and Das *(12)*	●			●	●	●	●	5
Jacobs *(13)*				●		●		2
Katabarwa, Mutabazi and Richards *(14)*	●	●		●		●		4
Katabarwa, Habomugisha and Agunyo *(15)*	●	●		●		●		4
Murayama, Taguchi and Murashima *(16)*	●	●		●		●	●	5
Sadruddin et al. *(17)*	●	●		●	●	●		5

Summary of qualitative findings

Objective	To identify, appraise, and synthesize qualitative research evidence on the use of community engagement strategies to promote CHW programme performance and utilization		
Perspective	Experiences and attitudes of stakeholders about lay health worker programmes in any country		
Included programmes	CHW programmes implemented in any country, with vulnerable populations		
Review finding	**Overall CERQual assessment of confidence**	**Explanation of CERQual assessment**	**Studies contributing to the review**
CHW performance			
CHW-level change			
CHW selection/nomination from and by community increases CHW commitment	Moderate confidence	This finding was graded as moderate confidence given the minor concerns about methodological limitations (potential for bias in recruitment strategies, as those CHWs interviewed were those who had stayed engaged in the programme, whereas the perspectives of those who had left were not incorporated) and adequacy (while five studies contributed to this review finding, this was not a primary outcome of interest in all five studies)	Abbey et al. *(3)*, Gopalan, Mohanty and Das *(12)*, Datiko et al. *(18)*, Frattaroli et al. *(19)*, Liverani et al. *(20)*
Political/power barriers to equitable community nomination of CHWs detract from CHW motivation	High confidence	This finding was graded as high confidence given only minor concerns about methodological limitations (in one study, the interviewer worked for the World Health Organization, which may have influenced respondent answers)	Najafizada, Labonté and Bourgeault *(21)*, Okuga et al. *(22)*, Saprii et al. *(23)*, Turinawe et al. *(24)*
Community engagement increases CHW retention	Low confidence	This finding was graded as low confidence given the minor concern about relevance, the moderate concern about methodological limitations (e.g., one study was conducted in a single site, with high levels of baseline community engagement, thus increasing the likelihood of ongoing CHW retention regardless of community engagement), and the major concern about adequacy (only two studies contributed to this review finding)	Frattaroli et al. *(19)*, Abimbola et al. *(25)*
Community engagement increases CHW job satisfaction	Moderate confidence	This finding was graded as moderate confidence given the moderate concern about methodological limitations (due to the high baseline of community engagement) and the minor concern about adequacy (only five studies contributed to this review; however, three provided rich quotations to support the view that community engagement increases CHW job satisfaction)	Datiko et al. *(18)*, Liverani et al. *(20)*, Abimbola et al. *(25)*, Razee et al. *(26)*, Wiggins et al. *(27)*
Outcomes: clients and community			

CHW selection/nomination from and by community increases CHW programme utilization	Moderate confidence	This finding was graded as moderate confidence given the moderate concern about methodological limitations (one study only included a single site, and one study included a potentially biased sample of CHWs who were present at the end of the study period, with the impetus to respond in a positive way) and moderate concern about adequacy (only four studies contributed to this outcome)	Frattaroli et al. *(19)*, Najafizada, Labonté and Bourgeault *(21)*, Okuga et al. *(22)*, Diakite, Keita and Mwebesa *(28)*
CHW selection/nomination from and by community increases community trust in CHWs	Moderate confidence	This finding was graded as moderate confidence given the minor concerns with methodological limitations (one study from a single site) and adequacy (eight studies contributed to this review; however, not all with rich data specific to the idea of CHW selection and community trust)	Datiko et al. *(18)*, Frattaroli et al. *(19)*, Najafizada, Labonté and Bourgeault *(21)*, Turinawe et al. *(24)*, De Jesus *(29)*, Mishra *(30)*, Singh, Cumming and Negin *(31)*, Srivastava et al. *(32)*
If CHWs are not from the target community or community selected, residing in the community increases CHW trust and utilization	Low confidence	This finding was graded as low confidence given the major concerns about adequacy (only one study contributed to this finding) and minor concerns about methodological limitations (interviews were brief and no member checking was conducted)	Zembe-Mkabile et al. *(33)*
Community engagement increases community trust in CHWs and programme utilization	High confidence	This finding was graded as high confidence given that there are no concerns, with the exception of a minor concern about methodological limitations (one study from a single site)	Datiko et al. *(18)*, Frattaroli et al. *(19)*, Liverani et al. *(20)*, Najafizada, Labonté and Bourgeault *(21)*, Wiggins et al. *(27)*, Diakite, Keita and Mwebesa *(28)*, De Jesus *(29)*, Mishra *(30)*, Singh, Cumming and Negin *(31)*, Srivastava et al. *(32)*, Hoy et al. *(34)*, Javanparast et al. *(35)*
Programme level			
Community actors			
Community engagement can support cultural competency in CHW programmes, which supports utilization and performance	High confidence	This finding was graded as high confidence given no concerns except for minor concerns about adequacy (six studies contributed to this review finding)	Najafizada, Labonté and Bourgeault *(21)*, Okuga et al. *(22)*, Razee et al. *(26)*, De Jesus *(29)*, Mishra *(30)*, Cook and Wills *(36)*
Community engagement increases community awareness of CHW programmes, leading to increased utilization	High confidence	This finding was graded as high confidence given no concerns except for minor concerns about methodological limitations (two of the studies provided inadequate details about methods to assess quality)	Elmardi et al. *(11)*, Liverani et al. *(20)*, De Jesus *(29)*, Mishra *(30)*, Hoy et al. *(34)*, George et al. *(37)*
Community engagement can increase community ownership/empowerment,	Moderate confidence	This finding was graded as moderate confidence given the moderate concern about methodological limitations (one study from a single site; another study conducted only in a neighbourhood with	Capell *(9)*, Gopalan, Mohanty and Das *(12)*, Frattaroli et al. *(19)*, Diakite, Keita and

leading to increased CHW programme uptake		successful CHW programme implementation, resulting in sample bias) and the minor concerns about coherence and adequacy (seven studies contributed to this review; however, not all studies provide rich data specific to community ownership/empowerment)	Mwebesa *(28)*, De Jesus *(29)*, Mishra *(30)*, George et al. *(37)*
Community engagement may increase community and local government financial contribution to CHW pay	Moderate confidence	This finding was graded as moderate confidence given the minor concern with methodological limitations (potentially biased sample as community with high level of community engagement at baseline) and moderate concern about adequacy (only two studies contributed to this review finding; however, they were very focused on remuneration)	Abbey et al. *(3)*, Abimbola et al. *(25)*
Community engagement in the case of stigmatized illnesses (e.g. HIV) may counterproductively lead the community to distance themselves from CHW programmes	Low confidence	This finding was graded as low confidence given the major concerns about adequacy (two studies contributing to this review finding) and minor concerns about methodological limitations (in one study, all interviewees chose to participate in the programme, reflective of a less vulnerable population)	Campbell et al. *(38)*, Masquillier et al. *(39)*
CHW community engagement and contributions may be undermined by overemphasis on "performance" and professional conduct, to neglect of CHW autonomy and informality	Low confidence	This finding was graded as low confidence given the major concerns about adequacy (only one study contributed to this finding)	Cook and Wills *(36)*
System level			
Community system			
Addressing gender/inequality in CHW programmes increases CHW performance	Moderate confidence	This finding was graded as moderate confidence given the minor concern about methodological limitations (potential bias in response due to interviewer) and moderate concern about adequacy (only three studies contributed to this finding)	Najafizada, Labonté and Bourgeault *(21)*, Saprii et al. *(23)*, Elazan et al. *(40)*
Community engagement strategies must take into account existing social hierarchies in communities to support CHW performance	Moderate confidence	This finding was graded as moderate confidence given the moderate concern about methodological limitations (concerns about sample bias in one study, and lack of clarity about CHW programme in another study) and moderate concern about adequacy (four studies contributed to this finding)	Okuga et al. *(22)*, Mishra *(30)*, Javanparast et al. *(35)*, George et al. *(37)*
Community engagement + CHW programme mitigates power imbalances, in turn increasing programme effectiveness	Moderate confidence	This finding was graded as moderate confidence given the moderate concern about methodological limitations (one study conducted in a single city; another study had lack of clarity about CHW programme) and the minor concern about adequacy (only three studies contributed to this finding; however, many data were specific to how community engagement mitigates power imbalances to increase programme effectiveness)	Frattaroli et al. *(19)*, George et al. *(37)*, Campbell and Mzaidume *(41)*

Community engagement may support CHWs in promoting community pride and combatting stigma among marginalized groups (e.g., female sex workers, people living with HIV)	Moderate confidence	This finding was graded as moderate confidence given the minor concerns about methodological limitations (in one study, all interviewees chose to participate in the programme, reflecting a less vulnerable population) and the moderate concern about adequacy (four studies contributed to this finding)	Hoy et al. *(34)*, Campbell et al. *(38)*, Masquillier et al. *(39)*, Campbell and Mzaidume *(41)*
The impact of community engagement on CHW performance is hampered by structural barriers (poverty, access to care, male domination, funding agencies)	Moderate confidence	This finding was graded as moderate confidence given the minor concerns about methodological limitations (two studies did not provide enough information to adequately assess study quality) and adequacy (eight studies contributed to this review finding, of which five provided rich data for synthesis)	Gopalan, Mohanty and Das *(12)*, Liverani et al. *(20)*, Saprii et al. *(23)*, Srivastava et al. *(32)*, Zembe-Mkabile et al. *(33)*, Javanparast et al. *(35)*, Elazan et al. *(40)*, Campbell and Mzaidume *(41)*, Cornish and Ghosh *(42)*
Health system			
Community engagement may promote local health sector involvement and strengthening	Moderate confidence	This finding was graded as moderate confidence given the minor concerns about methodological limitations (potential bias in sample selection) and adequacy (while nine studies contributed to this finding, some did not provide rich data for analysis, limiting their utility)	Datiko et al. *(18)*, Najafizada, Labonté and Bourgeault *(21)*, Okuga et al. *(22)*, Saprii et al. *(23)*, Mishra *(30)*, Zembe-Mkabile et al. *(33)*, Cook and Wills *(36)*, Campbell et al. *(38)*, Cornish and Ghosh *(42)*
Underresourced community/local health systems may undermine positive influence of community engagement on CHW performance	Moderate confidence	This finding was graded as moderate confidence given the minor concerns (sample section and social desirability bias in two studies) about methodological limitations and adequacy (four studies contributed to this finding)	Najafizada, Labonté and Bourgeault *(21)*, Saprii et al. *(23)*, Srivastava et al. *(32)*, Javanparast et al. *(35)*
Community engagement may be essential to enabling longer-term structural change	High confidence	This finding was graded as high confidence given the lack of concerns across all categories	Jacobs *(13)*, Liverani et al. *(20)*, Najafizada, Labonté and Bourgeault *(21)*, Wiggins et al. *(27)*, De Jesus *(29)*, George et al. *(37)*

Evidence to decision table

Recommendation 13

*WHO **recommends** the adoption of the following community engagement strategies in the context of practising CHW programmes:*

- pre-programme consultation with community leaders;
- community participation in CHW selection;
- monitoring of CHWs;
- selection and priority setting of CHW activities;
- support to community-based structures;
- involvement of community representatives in decision-making, problem solving, planning and budgeting processes.

Certainty of the evidence – moderate. Strength of the recommendation – strong.

Population: practising CHWs

Intervention: community engagement strategies vs not

Factors	Decision	Explanations/comments
Magnitude of desirable effects	How substantial are the desirable anticipated effects? ○ Trivial ○ Small ○ Moderate ● Large ○ Varies ○ Don't know	• Most quantitative, qualitative and mixed methods studies indicate that a range of community engagement strategies have beneficial impacts on CHW performance outputs, including CHW motivation, commitment, satisfaction and retention. Community engagement strategies also have beneficial impacts on CHW performance outcomes, including community trust of CHWs, and community awareness, support and sense of ownership of CHW programmes • Three RCTs indicate that community engagement strategies are effective in increasing CHW programme impact at the population level, all in the domain of maternal and child health outcomes among rural communities in low- and middle-income countries
Magnitude of undesirable effects	How substantial are the undesirable anticipated effects? ○ Large ○ Moderate ○ Small ○ Trivial ● Varies ○ Don't know	• There is some evidence on potential tension between communities and their leaders that may negatively impact the efforts of CHWs in community engagement strategies
Certainty of evidence	What is the overall certainty of the evidence of effects? ○ Very low ○ Low ● Moderate ○ High ○ No included studies	• The systematic review team rated the certainty of the evidence as moderate

Uncertainty or variability in how much people value the main outcomes	Is there important uncertainty about or variability in how much people value the main outcomes? ○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The systematic review team did not identify any studies assessing the values attached to the outcomes of interest • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	Does the balance between desirable and undesirable effects favour the intervention or the comparison? ○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ○ Probably favours the intervention ● Favours the intervention ○ Varies ○ Don't know	• The overwhelming majority of the evidence retrieved was supportive of the adoption of community engagement strategies
Resource use and cost-effectiveness	How large are the resource requirements (costs)? ○ Large costs ● Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• The systematic review team observed that the community engagement strategies reviewed included a broad range of activities, from presumably low cost (e.g., having a traditional elder choose CHWs) to moderate cost (e.g., ongoing outreach for the duration of the programme). Only one study, a very large-scale population intervention – Lewycka et al. *(2)* – specifically assessed the overall costs, which were deemed substantial, though this is not necessarily representative of the vast majority of studies included, for which the costs appear to be minimal to moderate. The same study found that the interventions examined were highly cost-effective
Impact on health equity	What would be the impact on health equity? ○ Reduced ○ Probably reduced ○ Probably no impact ○ Probably increased ● Increased ○ Varies ○ Don't know	• The systematic review team identified evidence suggesting that community engagement strategies support increased health equity. Data from large-scale CHW intervention trials indicate that community engagement strategies were associated with improved child and maternal health outcomes among vulnerable populations in low-income settings. Several descriptive and qualitative studies in the United States of America show beneficial effects of community engagement strategies in CHW programmes specifically designed for vulnerable populations (e.g., ethnic minorities, immigrants, poor and rural communities) that experience health disparities

Acceptability and feasibility of intervention	Is the intervention acceptable to key stakeholders and feasible? ○ No ○ Probably no ○ Probably yes ● Yes ○ Varies ○ Don't know	• The systematic review team identified scant research assessing the acceptability of community engagement strategies to a range of key stakeholders. Generally, descriptive studies and qualitative investigations indicate that community engagement strategies are acceptable to CHWs, to health care workers and professionals, and to local communities • A few studies suggest resistance by local leaders who fear their power is being usurped or threatened, some of whom interfere with fair processes of local CHW selection, instead choosing their favourites, who are often not recognized by the local communities • The stakeholder perception survey found that community engagement strategies had both high acceptability and feasibility

Annex 6.13 references

1. Bhutta ZA, Soofi S, Cousens S, Mohammad S, Memon ZA, Ali I et al. Improvement of perinatal and newborn care in rural Pakistan through community-based strategies: a cluster-randomised effectiveness trial. Lancet. 2011;377(9763):403–12.
2. Lewycka S, Mwansambo C, Rosato M, Kazembe P, Phiri T, Mganga A et al. Effect of women's groups and volunteer peer counselling on rates of mortality, morbidity, and health behaviours in mothers and children in rural Malawi (MaiMwana): a factorial, cluster-randomised controlled trial. Lancet. 2013;381(9879):1721–35.
3. Abbey M, Bartholomew LK, Nonvignon J, Chinbuah MA, Pappoe M, Gyapong M et al. Factors related to retention of community health workers in a trial on community-based management of fever in children under 5 years in the Dangme West district of Ghana. International Health. 2014;6:99–105.
4. Adams AM, Nababan HY, Hanifi SM. Building social networks for maternal and newborn health in poor urban settlements: a cross-sectional study in Bangladesh. PLoS ONE. 2015;10(4):e0123817.
5. Adejumo AO, Azuogu B, Okorie O, Lawal OM, Onazi OJ, Gidado M et al. Community referral for presumptive TB in Nigeria: a comparison of four models of active case finding. BMC Public Health. 2016;16:177.
6. Ahluwalia IB, Schmid T, Kouletio M, Kanenda O. An evaluation of a community-based approach to safe motherhood in northwestern Tanzania. International Journal of Gynaecology and Obstetrics. 2003;82(2):231–40.
7. Andersen MR, Yasui Y, Meischke H, Kuniyuki A, Etzioni R, Urban N. The effectiveness of mammography promotion by volunteers in rural communities. American Journal of Preventive Medicine. 2000;18(3):199–207.
8. Broadhead RS, Heckathorn DD, Weakliem DL, Anthony DL, Madray H, Mills RJ et al. Harnessing peer networks as an instrument for AIDS prevention: results from a peer-driven intervention. Public Health Report. 1998;113(Suppl. 1):42–57.
9. Capell JR. Evaluating a community participatory quality improvement process: what works and what can be improved? [dissertation]. Ann Arbor: ProQuest LLC; 2014.
10. Edward A, Ernst P, Taylor C, Becker S, Mazive E, Perry H. Examining the evidence of under-five mortality reduction in a community-based programme in Gaza, Mozambique. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2007;101(8):814–22.

11. Elmardi KA, Malik EM, Abdelgadir T, Ali SH, Elsyed AH, Mudather MA et al. Feasibility and acceptability of home-based management of malaria strategy adapted to Sudan's conditions using artemisinin-based combination therapy and rapid diagnostic test. Malaria Journal. 2009;8:39.

12. Gopalan SS, Mohanty S, Das A. Assessing community health workers' performance motivation: a mixed-methods approach on India's accredited social health activists (ASHA) programme. BMJ Open. 2012;2(5):e001557.

13. Jacobs L. The role of social capital in a community health worker model for grassroots advocacy [dissertation]. Tucson: University of Arizona; 2013.

14. Katabarwa NM, Mutabazi D, Richards FO Jr. Controlling onchocerciasis by community-directed, ivermectin-treatment programmes in Uganda: why do some communities succeed and others fail? Annals of Tropical Medicine and Parasitology. 2000;94(4):343–52.

15. Katabarwa MN, Habomugisha P, Agunyo S. Involvement and performance of women in community-directed treatment with ivermectin for onchocerciasis control in Rukungiri district, Uganda. Health and Social Care in the Community. 2002;10(5):382–93.

16. Murayama H, Taguchi A, Murashima S. Exploring the ideal combination of activity satisfaction and burden among health promotion volunteers: a cross-sectional study in Japan. BMC Public Health. 2013;13:205.

17. Sadruddin S, Khan IU, Bari A, Khan A, Ahmad I, Qazi SA. Effect of community mobilization on appropriate care seeking for pneumonia in Haripur, Pakistan. Journal of Global Health. 2015;5(1):010405.

18. Datiko DG, Yassin MA, Tulloch O, Asnake G, Tesema T, Jamal H et al. Exploring providers' perspectives of a community based TB approach in southern Ethiopia: implication for community based approaches. BMC Health Services Research. 2015;15:501.

19. Frattaroli S, Pollack KM, Jonsberg K, Croteau G, Rivera J, Mendel JS. Streetworkers, youth violence prevention, and peacemaking in Lowell, Massachusetts: lessons and voices from the community. Progress in Community Health Partnership. 2010;4(3):171–9.

20. Liverani M, Nguon C, Sok R, Kim D, Nou P, Nguon S et al. Improving access to health care amongst vulnerable populations: a qualitative study of village malaria workers in Kampot, Cambodia. BMC Health Services Research. 2017;17(1):335.

21. Najafizada SA, Labonté R, Bourgeault IL. Community health workers of Afghanistan: a qualitative study of a national program. Conflict and Health. 2014;8:26.

22. Okuga M, Kemigisa M, Namutamba S, Namazzi G, Waiswa P. Engaging community health workers in maternal and newborn care in eastern Uganda. Global Health Action. 2015;8(1):23968.

23. Saprii L, Richards E, Kokho P, Theobald S. Community health workers in rural India: analysing the opportunities and challenges accredited social health activists (ASHAs) face in realising their multiple roles. Human Resources for Health. 2015;13:95.

24. Turinawe EB, Rwemisisi JT, Musinguzi LK, de Groot M, Muhangi D, de Vries DH et al. Selection and performance of village health teams (VHTs) in Uganda: lessons from the natural helper model of health promotion. Human Resources for Health. 2015;13:73.

25. Abimbola S, Olanipekun T, Igbokwe U, Negin J, Jan S, Martiniuk A et al. How decentralisation influences the retention of primary health care workers in rural Nigeria. Global Health Action. 2015;8(1):26616.

26. Razee H, Whittaker M, Jayasuriya R, Yap L, Brentnall L. Listening to the rural health workers in Papua New Guinea: the social factors that influence their motivation to work. Social Science and Medicine. 2012;75(5):828–35.

27. Wiggins N, Johnson D, Avila M, Farquhar SA, Michael YL, Rios T et al. Using popular education for community empowerment: perspectives of community health workers in the Poder es Salud/Power for Health program. Critical Public Health. 2009;19(1):11–22.

28. Diakite O, Keita DR, Mwebesa W. Guinea. Village health committees drive family planning uptake: communities play lead role in increased acceptability, availability. CORE Group; 2009 (https://coregroup.org/wp-content/uploads/media-backup/documents/Case_Studies/Case_Study_VHC_Save_Guinea.pdf, accessed 28 July 2018).

29. De Jesus M. Mutuality at the center: health promotion with Cape Verdean immigrant women. Ethnicity and Health. 2009;14(1):45–59.

30. Mishra A. "Trust and teamwork matter": community health workers' experiences in integrated service delivery in India. Global Public Health. 2014;9(8):960–74.

31. Singh D, Cumming R, Negin J. Acceptability and trust of community health workers offering maternal and newborn health education in rural Uganda. Health Education Research. 2015;30(6):947–58.

32. Srivastava D, Prakash S, Adhish V, Nair K, Gupta S, Nandan D. A study of interface of ASHA with the community and the service providers in eastern Uttar Pradesh. Indian Journal of Public Health. 2009;53(3):133–6.

33. Zembe-Mkabile WZ, Jackson D, Sanders D, Besada D, Daniels K, Zamasiya T et al. The "community" in community case management of childhood illnesses in Malawi. Global Health Action. 2016;9(1):29177.

34. Hoy D, Southavilay K, Chanlivong N, Phimphachanh C, Douangphachanh V, Toole MJ. Building capacity and community resilience to HIV: a project designed, implemented, and evaluated by young Lao people. Global Public Health. 2008;3(1):47–61.

35. Javanparast S, Baum F, Labonte R, Sanders D. Community health workers' perspectives on their contribution to rural health and well-being in Iran. American Journal of Public Health. 2011;101(12):2287–92.

36. Cook T, Wills J. Engaging with marginalized communities: the experiences of London health trainers. Perspectives in Public Health. 2012;132(5):221–7.

37. George A, Blankenship KM, Biradavolu MR, Dhungana N, Tankasala N. Sex workers in HIV prevention: from social change agents to peer educators. Global Public Health. 2015;10(1):28–40.

38. Campbell C, Gibbs A, Maimane S, Nair Y. Hearing community voices: grassroots perceptions of an intervention to support health volunteers in South Africa. Journal of Social Aspects of HIV/AIDS. 2008;5(4):162–77.

39. Masquillier C, Wouters E, Mortelmans D, van Wyk B, Hausler H, Van Damme W. HIV/AIDS competent households: interaction between a health-enabling environment and community-based treatment adherence support for people living with HIV/AIDS in South Africa. PLoS ONE. 2016;11(3):e0151379.

40. Elazan SJ, Higgins-Steele AE, Fotso JC, Rosenthal MH, Rout D. Reproductive, maternal, newborn, and child health in the community: task-sharing between male and female health workers in an Indian rural context. Indian Journal of Community Medicine. 2016;41(1):34–8.

41. Campbell C, Mzaidume Z. Grassroots participation, peer education, and HIV prevention by sex workers in South Africa. American Journal of Public Health. 2001;91(12):1978–86.

42. Cornish F, Ghosh R. The necessary contradictions of “community-led” health promotion: a case study of HIV prevention in an Indian red light district. Social Science and Medicine. 2007;64(2):496–507.

6.14 Mobilization of community resources. In the context of CHW programmes, should practising CHWs mobilize wider community resources for health versus not?

Quantitative findings

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
Level of community participation in problem identification					
Engage community members by creating a plan to improve a community	NA	NA	1.31 1.0	145	⨁ [a]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			87 7 – 1.961		
Community impact: more community members voicing ideas or concerns about community issues	NA	NA	0.923;	145	⨁ [a]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			0.576 – 1.478		
Impact of community-engaged interventions on local health and social parameters					

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
Engage community members by identifying people or organizations that influence change	NA	NA	1.391; 0.755 – 2.55	145	⨁ [a]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			64		
Engage community members by helping someone to attend a public meeting	NA	NA	2.141; 1.407–3	145	⨁ [b]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			.208		
Engage community members by organizing an event	NA	NA	1.681; 1.140	145	⊕ [b]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			–2.478		
Community impact: increased participation in elections	NA	NA	1.850; 1.2	1455	⊕ [b]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			37–2.766		
Community impact: more community members attending public meetings	NA	NA	1.766; 1	145	⊕ [b]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			.158 – 2.693		
Community impact: community leaders taking action on an issue	NA	NA	1.816	145	⨁ [b]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			; 1.181 – 2.794		
Community impact: a concrete policy change	NA	NA	1.45	1	⨁ [b]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			7 1 ; 1 . 0 0 6 _ 2 . 1 5 0		
Sustainability of programme by local efforts					

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
Advocated a change to civic agency	NA	NA	1.474; 1.208 – 1.7	145	⨁ [b]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
				99	
Advocated a change to local government	NA	NA	1.665; 1.321–2	145	⨁ [b]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			.1		
Advocated a change to federal government	NA	NA	1.33 2; 1.143–1	145	⨁ [b]

Outcome(s)	Estimated risk (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Quality of evidence (GRADE)
	Control risk	Intervention risk			
			.55 (.51		

Notes	Notes on GRADE sores
⊕ indicates that the overall quality of evidence is very low. ⊕⊕ indicates that the overall quality of evidence is low. ⊕⊕⊕ indicates that the overall quality of evidence is moderate. ⊕⊕⊕⊕ indicates that the overall quality of evidence is high.	a. Not significant. b. CHWs in training intervention significantly more likely to report outcome than untrained CHWs.

GRADE quality assessment

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)
Level of community participation in problem identification	Outcomes	CHW-attributable changes in the community	Social cohesion	145	1: Ingram et al. *(1)*	36 months	●	No serious inconsistency[a]	Serious indirectness [1]	Serious imprecision[1]	NA[b]	No	No	No	⨁
Impact of community-engaged interventions on local health and social parameters	Outcomes	CHW-attributable changes in the community	Social cohesion	145	1: Ingram et al. *(1)*	36 months	●	No serious inconsistency[a]	Serious indirectness [1]	Serious imprecision[1]	NA[b]	No	No	No	⨁
Sustainability of programme by local efforts	Outcomes	CHW-attributable changes in the community	Social cohesion	145	1: Ingram et al. *(1)*	36 months	●	No serious inconsistency[a]	Serious indirectness [1]	Serious imprecision[1]	NA[b]	No	No	No	⨁

Notes

⨁ indicates that the overall quality of evidence is very low.

⨁⨁ indicates that the overall quality of evidence is low.

⨁⨁⨁ indicates that the overall quality of evidence is moderate.

⨁⨁⨁⨁ **indicates that the overall quality of evidence is high.**

Legend

● Low risk of bias.

● Unclear risk of bias.

● High risk of bias.

Notes on GRADE scores and sources

a. When only one study contributes evidence to an outcome, no serious inconsistency is assumed.

b. Undetected is recorded for publication bias, as it cannot be assessed for one study.

[1] **Ingram et al. *(1)*:** *serious indirectness*, **as the study was designed to measure the impact of a training programme on increasing community mobilization outcomes, not the impact of community mobilization on population health outcomes.** *Serious imprecision*, **due to an inadequate number of events (< 300) for the outcome. As this is a non-RCT, which starts out as low, this study receives a final GRADE assessment of** *very low* **after being rated down for high risk of bias, serious indirectness, and serious imprecision.**

Newcastle-Ottawa quality assessment for cross-sectional studies

Study	Selection				Comparability	Outcome		Total score
	Representativeness of sample	Sample size	Non-respondents	Ascertainment of exposure		Assessment of outcome	Appropriateness of statistical test	
Possible score	1	1	1	2	2	1	2	Max: 10
Ingram et al. *(1)*	✹	✹	✹	✹		✹	✹	6

Qualitative findings

Objective	To identify, appraise and synthesize qualitative research evidence examining whether practising CHWs should mobilize wider community resources for health		
Perspective	Programmes where CHWs leading activities have managed to mobilize wider community resources		
Included programmes	CHW programmes delivered in underserved or vulnerable populations among general population settings		
Review finding	**Overall CERQual assessment of confidence**	**Explanation of CERQual assessment**	**Studies contributing to the review**
Defining the role of CHWs as mobilizers of community change, and emphasizing mobilization activities as part of this role, enabled CHWs to advocate change, engage multiple stakeholders, and develop linkages with female sex workers	Moderate confidence; however, only one study included	Mobilization programmes can help to shift power to sex workers and enable them as change agents. However, this study does not provide evidence regarding whether mobilization approaches are more or less effective than traditional interventions	George et al. *(2)*

Evidence to decision table

Recommendation 14

*WHO **suggests** that CHWs contribute to mobilizing wider community resources for health by:*

- identifying priority health and social problems and developing and implementing corresponding action plans with the communities;
- mobilizing and helping coordinate relevant local resources representing different stakeholders, sectors and civil society organizations to address priority health problems;
- facilitating community participation in transparent evaluation and dissemination of routine community data and outcomes of interventions;
- strengthening linkages between the community and health facilities.

Certainty of the evidence – very low. Strength of the recommendation – conditional.

Population: practising CHWs

Intervention: mobilization of community resources vs not

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ● Large ○ Varies ○ Don't know	• Some limited evidence was found that equipping CHWs with a role and skills for mobilization of community resources makes them more effective in taking a change agent role • Despite the paucity of evidence, the GDG deemed the potential for these effects to be large
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ● Varies ○ Don't know	• No known or theoretical undesirable effects were identified
Certainty of evidence	● Very low ○ Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as very low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes

Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison ● Probably favours the intervention ○ Favours the intervention ○ Varies ○ Don't know	• The GDG was of the view that the benefits of CHW involvement in mobilization of community resources would justify support for this policy option, given the absence of known or theoretical undesirable effects
Resource use and cost-effectiveness	○ Large costs ○ Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ● Varies ○ Don't know	• The costs are likely to be of variable entity depending on which approaches to mobilize communities are adopted • However, the literature review identified no evidence on cost or cost-effectiveness
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The GDG was of the view that mobilization of vulnerable communities to be more proactive has the potential to improve health equity
Acceptability and feasibility of intervention	○ No ○ Probably no ○ Probably yes ● Yes ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing acceptability and feasibility • The stakeholder perception survey found the role of CHWs in community mobilization to be both acceptable and feasible

Annex 6.14 references

1. Ingram M, Chang J, Kunz S, Piper R, de Zapien JG, Strawder K. Women's health leadership to enhance community health workers as change agents. Health Promotion and Practice. 2016;17(3):391–9.

2. George A, Blankenship KM, Biradavolu MR, Dhungana N, Tankasala N. Sex workers in HIV prevention: from social change agents to peer educators. Global Public Health. 2015;10(1):28–40.

6.15 *Availability of supplies.* In the context of practising CHW programmes, what strategies should be used for ensuring adequate availability of commodities and consumable supplies over what other strategies?

Quantitative findings: summary of findings and overview of included studies

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
Intervention studies							
Chandani et al. *(1)*, Chandani et al. *(2)*, Shieshia et al. *(3)*	RCT	In Ethiopia, Malawi and Rwanda, 240, 139 and 321 CHWs were recruited respectively	Three intervention packages tailored to an mHealth intervention for supply chain management. These included product flow, data flow and effective people all summed up in an enhanced management (EM) intervention	Efficient product transport (EPT), which included product flow and data flow but not effective people	Mixed methods follow-up assessments were conducted	MHealth intervention over 18 months showed a steady increase: 79% to 99% in EM and 71% to 90% in EPT in routine use of cStock to report on stock levels A large majority (97% in EM and 91% in EPT) of health surveillance assistants reported that cStock had become their primary means for ordering or requesting health products from their resupply point Improved completeness of reports was reported for both groups (EM 85%, EPT 65%) Improved lead, a measure of responsiveness in the supply chain, was reported (EM group taking on average 12.8 days to fulfil an order and 26 days in the EPT group)	Service delivery, supply chain management
Observational studies							
Rowe et al. *(4)*	Cross-sectional	Siaya district, Kenya: 114 CHWs in an outpatient department of a hospital	Five quality improvement factors on CHW performance were assessed using two models	NA	Cross-sectional surveys were conducted on a sample of 192 ill child consultations	Adequacy of drug supplies for CHWs was not related to better guideline adherence (model 1: adjusted OR 1.74; 0.79–3.83, $P = 0.16$, vs model 2: adjusted OR 1.03; 0.50–2.12, $P = 0.94$) A lack of a relationship between using a flipchart job aid and guideline adherence was also noted (model 1 (no error): adjusted OR 3.04; 0.73–12.58, $P = 0.12$, vs model 2 (major error): adjusted OR 0.58; 0.18–1.95, $P = 0.38$)	Service delivery, supply chain management, competency

Study	Design	Setting and participants	Intervention	Comparison or control	Measures and data collection	Findings	Outcomes
Callaghan-Koru et al. *(5)*	Cross-sectional	Malawi: 29 health surveillance assistants were recruited	NA	NA	A mixed methodology of descriptive and qualitative components	One year after the training for CCM, 69% of the health surveillance assistants had all essential iCCM drugs in stock and 86% received a resupply in the last 3 months	Service delivery, supply chain management
Bagonza et al. *(6)*	Descriptive non-controlled	Wakiso district in Central Uganda: eligible CHWs (n = 300) from two randomly selected health subdistricts were interviewed	NA	NA	Information on CHW background characteristics, CHW prescription behaviours, health system support factors and availability of iCCM drugs. Multivariable logistic regression analysis was done to assess factors associated with availability of iCCM drugs	For drug availability, it was noted that only 8.3% of sites had a stock of all four drugs, with 11% of the CHWs not having access to any drugs at all Factors associated with iCCM drug availability were being supervised within the last month (adjusted OR 3.70; 95% CI, 1.22–11.24), appropriate drug prescriptions (more than 90%) (adjusted OR 3.71; 95 % CI, 1.38–9.96), regular submission of drug reports (adjusted OR 4.02; 95% CI, 1.62–10.10), and having a respiratory timer as a diagnostic tool (adjusted OR 3.11; 95% CI, 1.08–9.00)	Service delivery, supply chain management

Definitions

Odds ratio (OR). A measure of effect that is used to approximate relative risk (i.e., the likelihood that one group will experience the outcome given a certain exposure versus the likelihood that another group will experience the outcome given they were not exposed). When the OR is greater than 1.0, the risk is greater. When the OR is between 0 and 1, the risk is lower. When the risk is 1.0, there is no difference between groups. The further the OR is above or below 1.0, the larger the effect.

Correlation. A measure of association between two different constructs.

Significance or statistical significance. The probability that a finding was observed by chance alone. Traditionally, a finding is said to be "significant" when this probability is less than 0.05 (i.e., $P < 0.05$).

Confidence interval (CI). The estimated interval between which the measure of effect (e.g. the OR) would probably be observed if the study were conducted again on a similar sample of subjects.

Adjusted (e.g. adjusted OR). When an explanatory or causal factor's raw association with an outcome is statistically adjusted to take account of other potential explanatory factors.

GRADE quality assessment

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)
Medicine availability															
Identification of factors for iCCM drug availability	Output	CHW-attributable changes among individual clients	Medicine availability	303	1: Bagonza et al. *(6)*		●	No serious inconsistency[a]	No serious indirectness	Serious imprecision[1]	Undetected[b]	No	No	No	⊕
Drug availability one year after training + resupply in last 3 months	Output	CHW-attributable changes among individual clients	Medicine availability	29	1: Callaghan-Koru et al. *(5)*	1 year	●	No serious inconsistency[a]	No serious indirectness	Very serious imprecision[2]	Undetected[b]	No	No	No	⊕
Demand-based resupply via mHealth	Output	CHW-attributable changes among individual clients	Medicine availability	–	3: Chandani et al. *(1)*, Chandani et al. *(2)*, Shieshia et al. *(3)*	18 months	●	No serious inconsistency[a]	No serious indirectness	Serious imprecision[3]	Undetected[b]	No	No	No	⊕
Adherence to clinical guidelines according to two models	Output	CHW-attributable changes among individual clients	Medicine availability	114	1: Rowe et al. *(4)*		●	No serious inconsistency[a]	Very serious indirectness[4]	Serious imprecision[4]	Undetected[b]	No	No	No	⊕

Notes

⊕ indicates that the overall quality of evidence is very low.

⊕⊕ indicates that the overall quality of evidence is low.

⊕⊕⊕ indicates that the overall quality of evidence is moderate.

⊕⊕⊕⊕ indicates that the overall quality of evidence is high.

Legend

● Low risk of bias.

● Unclear risk of bias.

● High risk of bias.

Notes on GRADE scores and sources

a. When only one study contributes evidence to an outcome, no serious inconsistency is assumed.

b. Undetected is recorded for publication bias, as it cannot be assessed for one study.

Outcome(s)	Results	Classification	Measures	Participants (*n*)	Included studies (*n*)	Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Magnitude of effect	Dose–response gradient	Plausible control for confounding	Overall quality of evidence (GRADE)

[1] Bagonza et al. *(6)*: *serious imprecision*, due to an inadequate number of events (< 300) for the outcome. As this is a non-RCT, which starts out as low, this study receives a final GRADE assessment of *very low* after being rated down for serious imprecision.

[2] Callaghan-Koru et al. *(5)*: *very serious imprecision*, due a sample size < 50. As this is a non-RCT, which starts out as low, this study receives a final GRADE assessment of *very low* after being rated down for very serious imprecision.

[3] Chandani et al. *(1)*, Chandani et al. *(2)*, Shieshia et al. *(3)*: *serious imprecision*, due to an inadequate number of events (< 300) for the outcome or not stating the sample size. As these are non-RCTs, which start out as low, this study receives a final GRADE assessment of *very low* after being rated down for serious imprecision.

[4] Rowe et al. *(4)*: *very serious indirectness*, due to the cross-sectional survey that tested the association between having available drug supply and adhering to guidelines, rather than strategies to ensure adequate drug supply. *Serious imprecision*, due to an inadequate number of events (< 300) for the outcome. As this is a non-RCT, which starts out as low, this study receives a final GRADE assessment of *very low* after being rated down for very serious indirectness and serious imprecision.

Newcastle-Ottawa quality assessment for cross-sectional studies

Study	Selection				Comparability	Outcome		Total score
	Representativeness of sample	Sample size	Non-respondents	Ascertainment of exposure		Assessment of outcome	Appropriateness of statistical test	
Possible score	1	1	1	2	2	1	2	Max: 10
Rowe et al. *(4)*	✱	✱		✱	✱✱	✱	✱	7
Callaghan-Koru et al. *(5)*	✱	✱		✱	✱	✱	✱	6
Bagonza et al. *(6)*	✱	✱		✱✱	✱✱	✱	✱	8

Qualitative findings

Objective	To identify and synthesize qualitative data on strategies for ensuring adequate availability of commodities and consumable supplies		
Perspective	Experiences of stakeholders		
Included programmes	iCCM, mHealth		
Review finding	**Overall CERQual assessment of confidence**	**Explanation of CERQual assessment**	**Studies contributing to the review**
Transport difficulties exacerbate the challenge with drug stocks and availability	Moderate confidence	The finding is graded as moderate confidence, given the moderate concerns regarding relevance and adequacy	Ibrahim et al. *(7)*, Johnson et al. *(8)*
Team meetings improve coordination, solve problems, and improve relationships and engagement, but have their own challenges, such as finding resources and time to travel and attend meetings	Moderate confidence	The finding is graded as moderate confidence, given the moderate concerns regarding methodology, relevance and adequacy	Schiffman and Brinton *(9)*
Team meetings should have a predetermined agenda and use data for performance monitoring and identification of problems and solutions that are within the team's ability to address	Low confidence	The finding is graded as low confidence, given the moderate concerns regarding methodology, relevance, coherence and adequacy	Elliott et al. *(10)*
Provision of mobile phones to CHWs can improve supply chain management, as checks on supplies were consistent and replenishment was efficient, and transport costs were reduced	Moderate confidence	The finding is graded as moderate confidence, given the moderate concerns regarding relevance and adequacy	Rymkiewicz et al. *(11)*, Smikle et al. *(12)*

Evidence to decision table

Recommendation 15

*WHO **suggests** using the following strategies for ensuring adequate availability of commodities and consumable supplies, quality assurance, and appropriate storage, stocking and waste management in the context of CHW programmes:*

- integration in the overall health supply chain;
- adequate reporting, supervision, compensation, work environment management, appropriate training and feedback, and team quality improvement meetings;
- availability of mHealth to support different supply chain functions.

Certainty of the evidence – low. Strength of the recommendation – conditional.

Population: practising CHWs

Intervention: use of certain supply chain strategies vs other strategies.

Factors	Decision	Explanations/comments
Magnitude of desirable effects	○ Trivial ○ Small ○ Moderate ● Large ○ Varies ○ Don't know	• Several studies were identified presenting strategies associated with better supply chain management for CHWs
Magnitude of undesirable effects	○ Trivial ○ Small ○ Moderate ○ Large ○ Varies ● Don't know	• The included studies did not identify undesirable effects of the supply chain strategies analysed
Certainty of evidence	○ Very low ● Low ○ Moderate ○ High ○ No included studies	• The systematic review team assessed the overall certainty of the evidence as low
Uncertainty or variability in how much people value the main outcomes	○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability	• The studies included in the systematic review did not assess values and preferences on outcomes • The stakeholder perception survey identified coverage and quality of services, and competencies and motivation of CHWs, as the most important outcomes
Balance of benefits and harms	○ Favours the comparison ○ Probably favours the comparison ○ Does not favour either the intervention or the comparison	• A working supply chain system is an essential prerequisite for the functioning of CHW programmes

	○ Probably favours the intervention ● Favours the intervention ○ Varies ○ Don't know	
Resource use and cost-effectiveness	○ Large costs ● Moderate costs ○ Negligible costs and savings ○ Moderate savings ○ Large savings ○ Varies ○ Don't know	• Moderate costs are likely to be required for the programme • None of the studies included in the systematic review systematically measured costs and cost-effectiveness of applied strategies. However, cursory evidence included in studies indicates the mHealth strategy may help in reducing cost
Impact on health equity	○ Reduced ○ Probably reduced ○ Probably no impact ● Probably increased ○ Increased ○ Varies ○ Don't know	• The systematic review team did not identify any studies assessing the impact of the policy options on health equity. The GDG was of the view that strengthening CHW programmes through a more effective supply chain system would improve health equity
Acceptability and feasibility of intervention	○ No ○ Probably no ● Probably yes (acceptability) ● Yes (feasibility) ○ Varies ○ Don't know	• The stakeholder perception survey found various strategies for strengthening the supply chain for CHWs to be both acceptable and feasible, with the exception of the use of social media distribution aid, for which the acceptability and feasibility findings were more uncertain

Annex 6.15 references

1. Chandani Y, Andersson S, Heaton A, Noel M, Shieshia M, Mwirotsi A et al. Making products available among community health workers: evidence for improving community health supply chains from Ethiopia, Malawi, and Rwanda. Journal of Global Health. 2014;4(2):020405.
2. Chandani Y, Duffy M, Lamphere B, Noel M, Heaton A, Andersson S. Quality improvement practices to institutionalize supply chain best practices for iCCM: evidence from Rwanda and Malawi. Research in Social and Administrative Pharmacy. 2017;13(6):1095–109.
3. Shieshia M, Noel M, Andersson S, Felling B, Alva S, Agarwal S et al. Strengthening community health supply chain performance through an integrated approach: using mHealth technology and multilevel teams in Malawi. Journal of Global Health. 2014;4:2.
4. Rowe SY, Kelly JM, Olewe MA, Kleinbaum DG, McGowan JE Jr, McFarland DA et al. Effect of multiple interventions on community health workers' adherence to clinical guidelines in Siaya district, Kenya. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2007;101(2):188–202.
5. Callaghan-Koru JA, Hyder AA, George A, Gilroy KE, Nsona H, Mtimuni A et al. Health workers' and managers' perceptions of the integrated community case management program for childhood illness in Malawi: the importance of expanding access to child health services. American Journal of Tropical Medicine and Hygiene. 2012;87(Suppl. 5):61–8.

6. Bagonza J, Rutebemberwa E, Eckmanns T, Ekirapa-Kiracho E. What influences availability of medicines for the community management of childhood illnesses in central Uganda? Implications for scaling up the integrated community case management programme. BMC Public Health. 2015;15:1180.

7. Ibrahim FW, Schembri G, Taha H, Ariyanayagam S, Dhar J. Cervical surveillance in HIV-positive women: a genitourinary medicine clinic experience. Journal of Family Planning and Reproductive Health Care. 2009;35(2):101–3.

8. Johnson KB, Akwara P, Rutstein SO, Bernstein S. Fertility preferences and the need for contraception among women living with HIV: the basis for a joint action agenda. AIDS (London, England). 2009;23(Suppl. 1):S7–17.

9. Schiffman MH, Brinton LA. The epidemiology of cervical carcinogenesis. Cancer. 1995;76(10 Suppl.):1888–901.

10. Elliott G, Smith AC, Bensink ME, Brown C, Stewart C, Perry C et al. The feasibility of a community-based mobile telehealth screening service for Aboriginal and Torres Strait Islander children in Australia. Telemedicine Journal and e-Health: the official journal of the American Telemedicine Association. 2010;16(9):950–6.

11. Rymkiewicz PD, Heng YX, Vasudev A, Larbi A. The immune system in the aging human. Immunologic Research. 2012;53(1–3):235–50.

12. Smikle CB, Sorem KA, Satin AJ, Hankins GD. Physical and sexual abuse in a middle-class obstetric population. Southern Medical Journal. 1996;89(10):983–8.