

1 A new tool to measure approaches to supervision from the perspective of community health
2 workers: A prospective, longitudinal, validation study in seven countries.

3

4 Frédérique Vallières^{1*}, Philip Hyland^{1,2}, Eilish McAuliffe³, Ilias Mahmud^{4,5}, Olivia Tulloch⁶,
5 Polly Walker⁷, and Miriam Taegtmeier⁸

6

7 ¹Centre for Global Health, Trinity College Dublin, 7-9 Leinster Street South, Dublin 2, Ireland

8 ²School of Business, National College of Ireland, Mayor Street, IFSC, Dublin 1, Ireland. Philip

9 Hyland: philip.hyland@ncirl.ie

10 ³School of Nursing, Midwifery and Health Systems, University College Dublin, Dublin 4,

11 Ireland. Eilish McAuliffe: eilish.mcauliffe@ucd.ie

12 ⁴James P Grant School of Public Health, BRAC University, 68 ShahidTajuddin Ahmed

13 Sharani, Mohakhali, Dhaka 1212, Bangladesh. Ilias Mahmud: imahmudot@gmail.com

14 ⁵College of Public Health and Health Informatics, Qassim University, Bukayriah, Qassim, the

15 Kingdom of Saudi Arabia

16 ⁶Options Consultancy Service, St Magnus House, 3 Lower Thames Street, London EC3R 6HD,

17 United Kingdom. Olivia Tulloch: o.tulloch@options.co.uk

18 ⁷Global Centre for Health, HIV and WASH, World Vision International, Victoria Charities

19 Centre, 11 Belgrave Road, London, SW1V 1RB, United Kingdom. Polly Walker:

20 Polly.Walker@wvi.org

21 ⁸Department of International Public Health, Liverpool School of Tropical Medicine, Pembroke

22 Place, Liverpool, L3 5QA, United Kingdom. Miriam Taegtmeier:

23 Miriam.Taegtmeier@lstmed.ac.uk

24

25 *Corresponding author

26 E-mail: fvallier@ted.ie

27

Abstract

28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55

Background

The global scale-up of community health workers (CHWs) depends on supportive management and supervision of this expanding cadre. Existing tools fail to incorporate the perspective of the CHW (i.e. perceived supervision) in terms of supportive experiences with their supervisor. Aligned to the WHO's strategy on human resources for health, we developed and validated a simple tool to measure perceived supervision across seven low and middle-income countries.

Methods

Phase 1 was carried out with 327 CHWs in Sierra Leone. Twelve questions, informed by the extant literature on health worker supervision, were reduced to six questions using confirmatory factor analysis. Phase 2 employed structural equation modelling with 741 CHWs in six countries (Bangladesh, Ethiopia, Indonesia, Kenya, Malawi, Mozambique), to assess the factorial validity, predictive validity, and internal reliability of the questions at three time-points, over 8-months.

Results

We developed a robust, 6-item measure of perceived supervision (PSS), capturing regular contact, two-way communication, and joint problem-solving elements as being critical from the perspective of CHWs. When assessed across the six countries, over time, the PSS was also found to have good validity and internal reliability. PSS scores at baseline positively and significantly predicted a range of performance-related outcomes at follow-up.

Conclusion

The PSS is the first validated tool that measures supervisory experience from the perspective of CHWs and is applicable across multiple, culturally-distinct global health contexts with a wide range of CHW typologies. Simple, quick to administer, and freely available in eleven languages, the PSS could assist practitioners in the management of community health programmes.

56 **Keywords**

57 Community Health Workers, Supervision, Perceived Supervision Scale, Motivational

58 Outcomes, Scale Validation

59

60

Background

61 The important role of lower-cadre health workers in achieving Universal Health
62 Coverage (UHC) is widely recognised, with community health workers (CHWs)
63 frequently cited as a cost-effective, critical resource for the efficient delivery of primary
64 care in low- and middle-income contexts (LMICs) [1, 2]. Unfortunately, scaling up and
65 sustaining CHWs programme, as envisioned at Alma-Ata, has been challenging, with
66 wide variations in the availability, coordination, support and management of
67 community health worker programmes [3]. Accordingly, the most recent *Global
68 strategy on human resources for health: Workforce 2030* [4] published by the World
69 Health Organization (WHO) reiterates the need to harness the potential of community-
70 based health workers. Specifically, the strategy calls for a global effort to integrate
71 CHWs into national health-care systems as a means to improve their working
72 conditions, capacity, and motivation [4].

73 More recently, the WHO have also called for rigorous scientific research in the
74 area of community health workers to pay more attention to cross-cutting factors, such
75 as management and supervision, that enable community-based health worker
76 performance [5]. Decades of research on CHW initiatives to date have suggested
77 several cross-cutting factors that contribute to the success of CHW programmes [6].
78 Among these, supportive supervision consistently emerges as a key factor in
79 determining CHW performance, motivation, and retention [7].

80 In contrast to more ‘traditional’ methods of supervision which are frequently
81 characterised by performance audits, inspections, use of checklists, and controlling and
82 authoritarian attitudes [7-10], supportive supervision favours shared performance goals,
83 mentoring, and two-way communication [11]. Whereas traditional approaches are
84 frequently criticised for their failure to enhance health worker motivation [12-14],

85 supportive approaches to supervision have been shown to increase the impact of CHW
86 programmes as well as the productivity, motivation and job satisfaction of CHWs [7,
87 15-17]. Moreover, CHWs themselves express clear preferences for supportive
88 approaches that are responsive to the realities of the challenges they face in programme
89 implementation [14, 18].

90 In addition to supportive approaches to supervision, CHW programmes often
91 advocate for regular supervision of CHWs. Research suggests however that regular
92 interaction with one's supervisor is insufficient. When compared to colleagues who had
93 recently been supervised *and* felt supported by their supervisor, health workers who
94 had recently been supervised, but did *not* feel supported, were found to be less
95 productive [15]. This suggests that not only are health worker's perceptions of the
96 supervisory relationship significant, but that perceptions of the supportive nature of this
97 relationship is likely a more important predictor of work-related outcomes than
98 frequency alone. This view is consistent with well-established theories within the work
99 psychology literature, which state that subjective, cognitive appraisals of supervision
100 are critical factors in the prediction of a range of work performance-related factors (e.g.,
101 motivation, commitment, job satisfaction) [19].

102 While existing tools measure the supervision of CHWs (i.e. the "CHW
103 Assessment and Improvement Matrix" [20]) by assessing the frequency of supervision
104 and training of supervisors, these measures crucially ignore CHW perceptions of the
105 supervisory process and their impact on work-performance-related factors. Moreover,
106 such tools are lengthy, time-intensive, and require substantial programmatic input and
107 resources; all of which are at a premium within human resource for health programming
108 in LMICs. The need exists to develop a feasible, valid, and reliable measure of

109 perceived supervision that both recognises the experience of supervision from the
110 perspective of the individual health worker and that allows the CHW voice to be heard.

111 The current study aimed to develop and psychometrically validate a new, simple
112 measure of perceived supervision (the *Perceived Supervision Scale* (PSS)) that could
113 be used across multiple global health contexts. To maximise the utility of the PSS in
114 LMICs we sought to construct an easily-translatable measure, comprised of a limited
115 number of items that can be quickly and easily administered and scored; an approach
116 that should increase the likelihood of cross-cultural validity and subsequent use.

117 The development and validation of the PSS included two research phases. Phase
118 1, conducted in Sierra Leone, was exploratory and sought to determine the most
119 appropriate indicators of perceived supervision from an initial pool of test items. In
120 other words, we sought to determine which items, when included in a questionnaire,
121 measured perceived supervision among CHWs. Phase 2, conducted across six LMICs
122 and over a period of eight months, sought to provide a comprehensive assessment of
123 the psychometric properties of the PSS. Specifically, this phase assessed the predictive
124 validity, factorial validity, cross-cultural and temporal stability of the factor structure,
125 and the internal reliability of the PSS over time and across multiple cultural contexts.
126 In other words, we sought to determine whether the questionnaire, as developed in the
127 Sierra Leonean context also measured perceived supervision among CHWs across six
128 other contexts, and whether measures of perceived supervision using the PSS at
129 baseline, predicted a number of related human resource for health outcomes 8-months
130 later. Additionally, we assessed whether the total score on the PSS could be used by
131 implementers in the management and monitoring of CHW programmes.

132 **Methods**

133 *Participants and Procedures*

134 Phase 1 was conducted in Bonthe District, Sierra Leone among a convenience
135 sample of 327 CHWs, representing 98% of the CHWs active in the four chiefdoms of
136 Jong, Imperi, Sogbeni, and Kpanda Kemoh. Data collection took place over three weeks
137 in May 2012 as part of a longitudinal cohort study of CHWs participating in World
138 Vision Ireland’s Access to Infant and Maternal Health (AIM-Health) programme.
139 Phase 2 recruited a convenience sample of 741 CHWs from an additional six countries
140 (Bangladesh, Ethiopia, Kenya, Indonesia, Malawi and Mozambique) all of whom were
141 assessed across three time periods (baseline [T0], 4 months [T1], and 8 months [T2]).
142 CHWs were recruited in consultation with either national ministries of health
143 (Bangladesh, Malawi, Mozambique, Kenya), regional (Ethiopia) or district-level health
144 management teams (Indonesia), and based on the presence of a functioning CHW
145 programme in these districts. Data collection took place between October 2014 and
146 May 2015 as part of the REACHOUT research consortium
147 (www.reachoutconsortium.org). Demographic information for all participants is
148 reported in Table 1.

149

150 **INSERT TABLE 1 HERE**

151

152 *Development of the Initial Tool*

153 The 12 items of the PSS were initially constructed to capture aspects of
154 supervision described in the literature [21] [22]. Items are scored using a 5-point Likert
155 scale anchored by “strongly disagree” (1) and “strongly agree” (5). Items were designed
156 to capture key components of supervision, as identified from the literature, including
157 perceptions of regular contact (My supervisor meets with me regularly) and strong two-
158 way communication (My supervisor meets with me regularly to discuss problems and

159 solutions; My supervisor takes into consideration my views and ideas; and My
160 supervisor is a good communicator). These items were first translated in Phase 1 into
161 Krio, Sierra Leone's *lingua franca*. During Phase 2, the refined version of the PSS was
162 further translated into seven additional languages (Bangla, Kiswahili, Kamba, Bahasa-
163 Indonesia, Chichewa, Portuguese, and Amharic). Translated forms of the PSS are
164 available for free download at www.perceivedsupervisionscale.com. All versions were
165 piloted, revised, back-translated, and compared to the original English version prior to
166 being administered by trained enumerators. In the case of illiterate CHWs, the PSS was
167 administered with the help of an enumerator. In the case of literate CHWs, the PSS was
168 completed directly by the CHW. In both phases, enumerators were trained to administer
169 the PSS in the local languages and English.

170 In Phase 2, work-performance related factors were also assessed over time.
171 Adapted from Mbindyo et al. [23], the *Motivational Outcome Scale* is a 12 item, self-
172 report measure of work-performance related constructs: community commitment (2
173 items, $\alpha = .64$), organizational commitment (2 items, $\alpha = .44$), job satisfaction (4
174 items, $\alpha = .73$), and work conscientiousness (4 items, $\alpha = .73$). Each item was
175 assessed using a 5-point Likert Scale, anchored by "strongly disagree" (1) and "strongly
176 agree" (5). Among the current sample, the scale possessed satisfactory internal
177 reliability.

178 *Analysis*

179 During Phase 1, the initial pool of 12 PSS items were assessed using
180 confirmatory factor analysis (CFA) to develop a short, unidimensional measure of
181 perceived supervision (see Supplementary Table 2). CFA is a statistical technique that
182 tests whether items in a questionnaire effectively measure a theoretical construct, or
183 *latent construct*, that is itself not directly observable (i.e. perceived supervision) [24].

184 As Phase I was more exploratory in nature, we did not expect all 12 items to measure
185 perceived supervision in a consistent and robust manner. To determine which of these
186 12 items should be retained as the best measures of perceived supervision, we set an *a*
187 *priori* criterion for item retention whereby only items with factor loadings¹ >.55
188 (equalling 30% of variance explained by the latent variable) were retained [25]. In
189 addition to consulting factor loadings, we also consulted modification indices produced
190 in Mplus (Version 7.4). Modification indices provided suggestions of additional items
191 that could be removed to improve model fit (i.e. items with covarying residuals) [26].

192 Phase 2 also used CFA procedures to determine the factorial validity of the PSS.
193 In addition, structural equation modelling (SEM) methods were used to assess whether
194 perceived supervision scores, as measured by the PSS at baseline (Time 0), predicted
195 the four criterion variables of the Motivational Outcomes Scale at endline (Time 2),
196 controlling for sex and educational status. Here, SEM was chosen to assess the
197 predictive validity of the PSS as it allows for all effects in the model to be estimated
198 simultaneously. In other words, SEM methods were used to test whether the
199 administration of the PSS scale at earlier stages of CHW programmes predicted a range
200 of meaningful human resource for health-related outcomes throughout later stages of a
201 CHW programme, whereby job satisfaction, organizational commitment, community
202 commitment, and work conscientiousness were measured as known determinants of
203 CHW programme success. The internal reliability of the PSS was assessed using
204 composite reliability analysis [27], and descriptive statistics were calculated for each
205 country and at each assessment period.

¹ Depicted as λ_{pre} in Supplementary Table 2, factor loadings indicate what proportion of the variance in each item on the questionnaire can be explained by the underlying latent construct.

206 Analyses were conducted in Mplus 7.4 [28] using the mean and variance-
207 adjusted weighted least squares (WLSMV) estimator. The WLSMV estimator provides
208 accurate parameter estimates, standard errors, and test-statistics when ordinal indicators
209 are used [29]. Missing data was managed using the default pairwise present analysis
210 method. Standard recommendations for assessing the fit of the CFA and SEM models
211 were followed [30] whereby a non-significant chi-square (χ^2) result indicates good
212 model fit; Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) values $> .90$
213 indicate good fit; Root-Mean-Square Error of Approximation (RMSEA) with 90%
214 confidence interval (RMSEA 90% CI) values $< .08$ reflect good fit; and values < 1.0
215 for the Weighted Root Mean Square Residual (WRMR) indicate good model fit. In
216 other words, models that met these criteria were seen to be a ‘good’ representation of
217 perceived supervision.

218

Results

219 Phase 1: Development of the Perceived Supervision Scale

220 The fit of the unidimensional, 12-item model to the sample data was poor ($\chi^2 =$
221 355.417, $df = 54$, $p < .001$; CFI = .757; TLI = .703; RMSEA [90% CI] = .131 [.119-
222 .145]; WRMR = 1.739). Inspection of the model parameters indicated that six items
223 failed to reach the *a priori* criterion of factor loadings $> .55$ on the Perceived
224 Supervision factor (Supplementary Table 2). The unidimensional model was
225 subsequently re-estimated based on the remaining six items and model fit was
226 acceptable ($\chi^2 = 43.952$, $df = 9$, $p < .001$; CFI = .961; TLI = .934; RMSEA [90% CI] =
227 .110 [.079-.143]; WRMR = .910). The factor loadings for the six items were all positive,
228 statistically significant, and of a robust magnitude.

229 Phase 2: Validity of the Perceived Supervision Scale

230 Table 2 reports the CFA results for the six-item, unidimensional model of the
231 PSS across six nations, and at three assessment periods. In most cases the χ^2 values
232 were statistically significant and the RMSEA values were above the suggested cut-off
233 point of .08. However, rejection of the models based on these indices is not warranted
234 given the tendency for the χ^2 to generate Type 1 errors, and the RMSEA to generate
235 Type 2 errors in models with few degrees of freedom [31]. Contrastingly, the CFI, TLI,
236 and WRMR results provided consistent support for the factorial validity of the PSS. In
237 all 17 assessments, the CFI, TLI, and WRMR results satisfied the criteria for excellent
238 model fit. Overall, the CFA results provide support for the validity of a unidimensional
239 structure of the PSS that is stable over time, and cross-culturally consistent.

240 **INSERT TABLE 2 HERE**

241 The standardised factor loadings for the PSS across each nation, at each assessment,
242 are reported in Supplementary Table 3. Factor loadings at T0 were all positive,
243 significant ($p < .001$), and robust, with mean factor loadings ranging from .68
244 (Indonesia) to .92 (Kenya). Similarly, at T1 all factor loadings were positive, significant
245 ($p < .001$), and robust, with mean factor loadings ranging from .74 (Indonesia) to .83
246 (Ethiopia). At T2, there was greater variability in the performance of the model
247 parameters. Within the Indonesian sample it was necessary to add a residual covariance
248 between two items with the lowest factor loadings (PSS4 and PSS6: factor loadings $<$
249 .50) to achieve acceptable model fit. Additionally, within the Ethiopian sample two
250 items possessed weak factor loadings (PSS2 = .11 and PSS4 = .22). Nonetheless, mean
251 factor loadings were generally robust, ranging from .50 (Ethiopia) to .91 (Bangladesh).

252 Given the stability of the unidimensional structure of the PSS across nations,
253 and time, all PSS data at T0 was merged. Model fit of this consolidated data was
254 satisfactory ($N = 710$; $\chi^2 = 138.936$, $df = 9$, $p < .001$; CFI = .987; TLI = .979; RMSEA

255 [90% CI] = .143 [.122-.164]; WRMR = .864), and therefore used to assess predictive
256 validity².

257 *Predictive Validity of the Perceived Supervision Scale*

258 A PSS latent variable modelled at T0 was used to predict the summed scores of
259 four criterion variables (job satisfaction, organizational commitment, community
260 commitment, and work conscientiousness) measured eight months later (T2),
261 controlling for sex and educational status. The fit of the model to the data was excellent
262 ($\chi^2 = 91.276$, $df = 41$, $p < .001$; CFI = .991; TLI = .986; RMSEA [90% CI] = .045 [.033-
263 .058]; WRMR = .847). As detailed in Table 3, the model explained between 5.8% and
264 16.4% of variance in each of the criterion variables, and perceived supervision
265 positively predicted all variables (β values ranged from .16 to .30).

266

267 **INSERT TABLE 3 HERE**

268

269

270 *Internal reliability and descriptive statistics for the PSS*

271 Composite reliability analyses indicated that the PSS possesses satisfactory
272 internal reliability (Supplementary Table 3), indicating that the six items were internally
273 consistent and serve as accurate measures of perceived supervision. In every national
274 context, and at each assessment period, the reliabilities ranged from .68 to .97.
275 Descriptive statistics for the PSS across all nations, at each assessment period, are
276 presented in Table 4.

277 **INSERT TABLE 4 HERE**

278

² A unidimensional model indicates that the PSS should be scored by summing questions PSS1-PSS6 to produce a total PSS score.

279

280

Discussion

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

The Perceived Supervision Scale is the first validated tool developed for collecting CHW perceptions of their supervision. The tool is brief, robust and can be applied across multiple, culturally-distinct global health contexts with a wide range of CHW typologies. Despite its recognised importance of supervision in CHW programming, supervision is often one of the weakest and most difficult elements of CHW programming to implement consistently [9, 32]. The factor structure of the PSS allows researchers and implementers to calculate a sum score of perceived supervision within CHW programming. Specifically, the total PSS score allows for a greater understanding the nature of a positive supervisory relationship. Furthermore, it grants the ability to managers to detect problematic supervisory interactions, prompt the introduction of stronger training programmes, and where necessary, the reorganisation of supervisory arrangements, contributing to the sustainability of CHW programmes. The ability for CHW programme managers to monitor the interpersonal supervisory relationships of CHWs could help prevent deleterious work performance outcomes associated with high staff turnover and loss of worker motivation [7, 33]. The development of the PSS therefore represents a valuable contribution to global efforts to address human resource for health shortages and towards achieving UHC. Furthermore, the development of the PSS contributes towards addressing more recent calls for rigorous approaches towards scale development for human resource for health programming [34].

Phase 1 served to derive the most appropriate indicators of perceived supervision. From an initial pool of 12 item statements, developed from the extant literature on CHW supervision, six items were retained. Consistent with previous

304 literature, the items retained as part of the final PSS, reflect the importance of *both*
305 supportive and regular aspects of supervision. Interestingly, those items associated with
306 more traditional forms of supervision (i.e. controlling or negative interactions), were
307 least reflective of the nature of perceived supervision among this sample of CHWs.
308 This suggests that CHWs in Sierra Leone perceived the supervision process as a
309 generally positive, supportive, and regular experience. The items retained as part of the
310 supportive supervision factor offer additional insight into what content or skills should
311 be emphasised or included as part of supervision training programmes. More
312 specifically, the items retained in the PSS are consistent with evidence that a supportive
313 supervisor should: meet regularly with CHWs, offer opportunities for knowledge
314 sharing and refresher training [33], recognise and appreciate the work and efforts of a
315 CHW, take into account the views and ideas of CHWs, and communicate effectively
316 with the CHW [11].

317 As it was possible that the observed findings from Phase 1 reflected the
318 idiosyncratic responses of the Sierra Leonean CHWs, it was imperative to assess the
319 replicability of these findings in alternate contexts. Phase 2 confirmed the PSS's
320 unidimensional structure across multiple samples of CHWs from different contexts,
321 cadres, cultures, and demographics. Additionally, the factorial validity of the PSS was
322 evidenced across time, with the scale exhibiting stable psychometric properties
323 (reliability and validity) over a period of eight months. Furthermore, the PSS positively
324 predicted a range of work-performance related indicators eight months later including
325 job satisfaction, work conscientiousness, community commitment, and organizational
326 commitment, while controlling for sex and education. These results indicate that CHWs
327 who perceive greater levels of supervision (i.e. supportive) report greater job
328 satisfaction, work conscientiousness and higher levels of both community and

354 Simple and quick to administer, and currently available in nine languages, the validated
355 PSS has the potential to contribute towards a more accurate understanding of CHW's
356 perspectives of supervision, as a critical determinant of successful CHW programmes
357 across a wide range of contexts.
358

359 **Declarations**

360 **Ethics Approval and Consent to Participate**

361 Ethical approval was obtained from the Ethical Review Committee, Trinity College
362 Dublin and Connaught Hospital Freetown (Phase 1), in addition to the Liverpool School
363 of Tropical Medicine Research Ethics Committee and six local ethics committees
364 (Phase 2) (Supplementary Table 1).

365 **Consent for Publication**

366 Not Applicable

367 **Availability of Data and Material**

368 The datasets used and/or analysed during the current study are available from the
369 corresponding author on reasonable request.

370 **Competing Interests**

371 The authors declare that they have no competing interests

372 **Funding**

373 Phase 1 was funded by Irish Aid through World Vision Ireland's Access to Infant and
374 Maternal (AIM-Health) programme, as well as the Department for International
375 Development and the United Kingdom through their Programmes Partnership
376 Agreement. Phase 2 of this research was nested within REACHOUT, funded by the
377 European Union FP7 grant (number 306090).

378

379 **Author Contributions**

380 FV conducted the literature review, designed the study and, together with EM, designed
381 the items and coordinated Phase 1 data collection and analysis. PH and FV led the
382 overall data analysis. FV, PH, PW and MT contributed equally to the manuscript
383 writing. MT, OT, and IM coordinated the data collection and extracted the data for

384 Phase 2. All authors contributed to the critical interpretation of the results and approved
385 the final version. FV is the final guarantor of the manuscript.

386

387 **Acknowledgements**

388 The authors would like to thank the CHWs in seven countries who gave their
389 time to answers questions about their perceptions of supervision. Phase I of this
390 research was made possible thanks to the support of Irish Aid and the people of Ireland
391 through World Vision Ireland's Access to Infant and Maternal (AIM-Health)
392 programme as well as the Department for International Development and the United
393 Kingdom through their Programmes Partnership Agreement. A special thanks to Allieu
394 Bangura and Joseph Musa of World Vision Sierra Lone for their assistance with
395 organising the collection of the data.

396 Phase 2 of this research was nested within REACHOUT, a multi-country
397 research consortium aiming to maximize the equity, effectiveness and efficiency of
398 close-to-community services in rural areas and urban slums. Robinson Karuga,
399 Kingsley Chikaphupha, Daniel Gemechu, RaliciaLimato, Irin Akhter and
400 SozinhoNdima from the country REACHOUT teams organised the collection, cleaning
401 and analysis of country data. The REACHOUT Consortium is funded by the European
402 Union FP7 grant (number 306090). This document reflects only the authors' views and
403 the European Union is not liable for any use that may be made of the information
404 contained therein.

405

406 **References**

- 407 1. Gilmore B, McAuliffe E: **Effectiveness of community health workers delivering**
408 **preventive interventions for maternal and child health in low- and middle-**
409 **income countries: a systematic review.** *BMC Public Health* 2013, **13**:847.
- 410 2. McPake B, Edoke I, Witter S, Kielmann K, Taegtmeyer M, Dieleman M, Vaughan K,
411 Gama E, Kok M, Datiko D *et al*: **Cost-effectiveness of community-based**
412 **practitioner programmes in Ethiopia, Indonesia and Kenya.** *Bulletin of the World*
413 *Health Organization* 2015, **93**(9):631-639a.
- 414 3. Tulenko K, Mogedal S, Afzal MM, Frymus D, Oshin A, Pate M, Quain E, Pinel A,
415 Wynd S, Zodpey S: **Community health workers for universal health-care**
416 **coverage: from fragmentation to synergy.** *Bulletin of the World Health*
417 *Organization* 2013, **91**(11):847-852.
- 418 4. World Health Organization: **Global strategy on human resources for health:**
419 **workforce 2030.** . In. Geneva, Switzerland: World Health Organization,; 2016.
- 420 5. Maher D, Cometto G: **Research on community-based health workers is needed to**
421 **achieve the sustainable development goals.** *Bulletin of the World Health*
422 *Organization* 2016, **94**(11):786.
- 423 6. Kok MC, Dieleman M, Taegtmeyer M, Broerse JE, Kane SS, Ormel H, Tijm MM, de
424 Koning KA: **Which intervention design factors influence performance of**
425 **community health workers in low- and middle-income countries? A systematic**
426 **review.** *Health policy and planning* 2014.
- 427 7. Hill Z, Dumbaugh M, Benton L, Kallander K, Strachan D, ten Asbroek A,
428 Tibenderana J, Kirkwood B, Meek S: **Supervising community health workers in**
429 **low-income countries--a review of impact and implementation issues.** *Glob*
430 *Health Action* 2014, **7**:24085.
- 431 8. Bosch-Capblanch X, Garner P: **Primary health care supervision in developing**
432 **countries.** *Trop Med Int Health* 2008, **13**(3):369-383.

- 433 9. Crigler L, Gergen J, Perry H: **Supervision of Community Health Workers**. In:
434 *Developing and Strengthening Community Health Worker Programs at Scale: A*
435 *Reference Guide for Program Managers and Policy Makers*. Baltimore (MD): USAID
436 and MCHIP; 2014: 10.11-10.26.
- 437 10. Hernandez AR, Hurtig AK, Dahlblom K, San Sebastian M: **More than a checklist: a**
438 **realist evaluation of supervision of mid-level health workers in rural Guatemala**.
439 *BMC Health Serv Res* 2014, **14**:112.
- 440 11. Marquez L, Kean L: **Making supervision supportive and sustainable: new**
441 **approaches to old problems**. In. Washington DC: USAID: Maximizing Access and
442 Quality (MAQ) Initiative; 2002.
- 443 12. Kok M, Muula A: **Motivation and job satisfaction of health surveillance assistants**
444 **in Mwanza, Malawi: an explorative study**. *Malawi Med J* 2013, **25**:5-11.
- 445 13. Jaskiewicz W, Tulenko K: **Increasing community health worker productivity and**
446 **effectiveness: a review of the influence of the work environment**. *Human*
447 *resources for health* 2012, **10**(1):38.
- 448 14. Ndimba SD, Sidat M, Give C, Ormel H, Kok MC, Taegtmeier M: **Supervision of**
449 **community health workers in Mozambique: a qualitative study of factors**
450 **influencing motivation and programme implementation**. *Human resources for*
451 *health* 2015, **13**:63.
- 452 15. Frimpong JA, Hellinginger S, Awoonor-Williams JK, Yeji F, Phillips JF: **Does**
453 **supervision improve health worker productivity? Evidence from the Upper East**
454 **Region of Ghana**. *Trop Med Int Health* 2011.
- 455 16. Willis-Shattuck M, Bidwell P, Thomas S, Wyness L, Blaauw D, Ditlopo P:
456 **Motivation and retention of health workers in developing countries: a systematic**
457 **review**. *BMC Health Services Research* 2008, **8**(1):247.
- 458 17. McAuliffe E, Daly M, Kamwendo F, Masanja H, Sidat M, de Pinho H: **The critical**
459 **role of supervision in retaining staff in obstetric services: a three country study**.
460 *PLoS One* 2013, **8**(3):e58415.

- 461 18. Kok MC, Vallières F, Tulloch O, Kumar MB, Kea AZ, Karuga R, Ndimba SD,
462 Chikaphupha K, Theobald S, Taegtmeier M: **Does supportive supervision enhance**
463 **community health worker motivation? A mixed-methods study in four African**
464 **countries.** *Health policy and planning* 2018;czy082-czy082.
- 465 19. Deci EL, Flaste R: **Why we do what we do: Understanding Self-Motivation.** New
466 York: Penguin Group; 1995.
- 467 20. Crigler L, Furth R, Bjerregaard D: **CommunitH Health Worker Assessment and**
468 **Improvement Matrix (CHW AIM): A Toolkit for Improving Community Health**
469 **Worker Programs and Services.** In. Bethesda, MD: University Reserach Co., LLC;
470 2011.
- 471 21. Mathauer I, Imhoff I: **Health worker motivation in Africa: the role of non-**
472 **financial incentives and human resource management tools.** *Human resources for*
473 *health* 2006, **4**:24.
- 474 22. May DR, Gilson RL, Harter LM: **The psychological conditions of meaningfulness,**
475 **safety and availability and the engagement of the human spirit at work.** *Journal*
476 *of Occupational and Organizational Psychology* 2004, **77**:11-37.
- 477 23. Mbindyo PM, Blaauw D, Gilson L, English M: **Developing a tool to measure health**
478 **worker motivation in district hospitals in Kenya.** *Human resources for health*
479 2009, **7**:40.
- 480 24. Thompson B: **Exploratory and confirmatory factor analysis.** Washington, DC:
481 American Psychological Association; 2004.
- 482 25. Comrey AL, Lee HB: **A first course in factor analysis.** Hillsdale, NJ: Erlbaum
483 Associates.; 1992.
- 484 26. Byrne BM: **Structural Equation Modeling with Mplus: Basic concepts,**
485 **Applications, and Programming.** Hove, East Sussex: Routledge Taylor & Francis
486 Group; 2012.
- 487 27. Raykov T: **Behavioral Scale Reliability and Measurement Invariance Evaluation**
488 **Using Latent Variable Modeling.** *Behavior Therapy* 2004, **35**:299-331.

- 489 28. Muthen LK, Muthen BO: **MPlus Statistical Analysis with Latent Variables**. In.,
490 7.4 edn. Los Angeles: Muthen and Muthen; 2013.
- 491 29. Flora DB, Curran PJ: **An empirical evaluation of alternative methods of**
492 **estimation for confirmatory factor analysis with ordinal data**. *Psychol Methods*
493 2004, **9**(4):466-491.
- 494 30. Hu L, Bentler PM: **Cutoff criteria for fit indices in covariance structure analysis:**
495 **Conventional criteria versus new alternatives**. *Structural Equation Modeling* 1999,
496 **6**:1-55.
- 497 31. Kenny DA, Kaniskan B, McCoach DB: **The performance of RMSEA in models**
498 **with small degrees of freedom**. *Sociological Methods & Research* 2015, **44**:486-
499 507.
- 500 32. Lehmann U, Sanders D: **Community Health Workers: What do we know about**
501 **them? The state of the evidence on programmes, activities, costs and impact on**
502 **health outcomes of using community health workers**. In. Geneva: World Health
503 Organization; 2007.
- 504 33. Strachan DL, Kallander K, Ten Asbroek AH, Kirkwood B, Meek SR, Benton L,
505 Conteh L, Tibenderana J, Hill Z: **Interventions to Improve Motivation and**
506 **Retention of Community Health Workers Delivering Integrated Community**
507 **Case Management (iCCM): Stakeholder Perceptions and Priorities**. *Am J Trop*
508 *Med Hyg* 2012, **87**(5 Suppl):111-119.
- 509 34. Borghi J, Lohmann J, Dale E, Meheus F, Goudge J, Oboirien K, Kuwawenaruwa A:
510 **How to do (or not to do)... Measuring health worker motivation in surveys in**
511 **low- and middle-income countries**. *Health policy and planning* 2017.
- 512 35. Wolf EJ, Harrington KM, Clark SL, Miller MW: **Sample Size Requirements for**
513 **Structural Equation Models: An Evaluation of Power, Bias, and Solution**
514 **Propriety**. *Educ Psychol Meas* 2013, **76**(6):913-934.
- 515