

2 **Measuring the strength of maternal, newborn and child health care**
3 **implementation and its association with childhood mortality risk in**
4 **three rural districts of Tanzania.**

5
6 **Short title:** Does the strength of maternal, newborn and child health care implementation affect
7 child survival?

8
9 Colin Baynes^{1*}, Almamy Malick Kanté², Amon Exavery³, Tani Kassimu⁴, Gloria Sikustahili⁴,
10 Hildegalda Mushi⁴, Kate Ramsey⁵, Kenneth Sherr¹, Bryan Weiner¹, James F. Phillips⁶

11
12
13 ¹Department of Global Health, University of Washington, Seattle, WA, USA

14 ²Department of International Health, Johns Hopkins University, Baltimore, MD, USA

15 ³Pact International, Dar es Salaam, Tanzania

16 ⁴Ifakara Health Institute, Dar es Salaam, Tanzania

17 ⁵Scope Impact, Brooklyn, NY, USA

18 ⁶Department of Population and Family Health, Columbia University, New York, NY, USA

19
20
21 *Corresponding author

22 Email: colin.baynes@gmail.com

23
24
25 CB conceptualized the study, had management and coordination responsibility for the data
26 collection, developed the analytic methodology, led the analysis of data, and wrote the original
27 and final draft of the paper. AMK and KR had oversight and leadership responsibility for the
28 research activity, reviewed and edited the paper. AE and TK curated the data used for this
29 analysis, participated in data collection, and reviewed and edited the paper. GS and HM
30 participated in data collection, and reviewed and edited the paper. KS and BW validated the
31 reproducibility of the results of the analysis, reviewed and edited the paper. JFP acquired
32 financial support for this study, validated the reproducibility of the results of the analysis,
33 reviewed and edited the paper.

34
35

36

37 **Abstract**

38 This observational cohort study explores the association between maternal, newborn and child
39 health care implementation strength and child survival in rural Tanzania from 2011-2015. We used
40 data from a 2011 service availability and readiness assessment that quantified primary health care
41 facilities' ability to implement maternal, newborn and child health services and a population-level
42 household survey that measured the utilization of such services to develop domain-specific
43 summary measures of the effective coverage of facility-based maternal, newborn and child health
44 care. We reduced domain specific effective coverage scores into fewer, independent scales of
45 implementation strength using principal components analysis, and integrated them into gradients
46 of the collective implementation strength exerted by groups of facilities on villages they served
47 using Bayesian mixed effects models. We linked these scales to longitudinal data on the survival
48 of children that were born in the catchment areas of the surveyed health facilities during the
49 assessment period and followed up until December 31, 2015. We fit survival time models to
50 estimate the relationship between implementation strength and child mortality. Increases in the
51 implementation strength gauged by our first scale, which represented general facility readiness and
52 the provision of antenatal, postnatal, and early childhood preventive services, were associated with
53 child mortality risks that were, on average, 0.62 times lower. Increases in implementation strength
54 gauged by our second scale, which represented sick childcare service provision, were associated
55 with child mortality risks that were, on average, 0.56 times lower. We detected no significant child
56 mortality response to our third scale, which represented intrapartum care provision. The findings
57 suggest that strong implementation of antenatal, postnatal, early childhood preventive services and
58 sick child care can accelerate child mortality reduction and that routine data on service availability
59 and readiness can be used to measure health systems strengthening and its impacts.

60 **Introduction**

61 In the past two decades, the spread of evidence-based child survival interventions has precipitated
62 large-scale reductions in child mortality globally; however, considerable geographic disparities
63 exist (1,2). Preventable loss of life during childhood remains concentrated in sub-Saharan Africa
64 (sSA), where children are between six and fifteen times more likely to die before reaching age five
65 compared to children in more developed regions (3,4). To address this inequality, health sectors in
66 sSA have departed disease-specific programming strategies and adopted holistic frameworks for
67 bundling low cost, evidence-based interventions (EBI) along the maternal, newborn and child
68 health (MNCH) ‘continuum of care’, and integrated these bundles into primary health care (PHC)
69 delivery systems (5–8). This introduced challenges to the task of assessing the impact of child
70 survival programs. Whereas evaluations had traditionally focused on the effectiveness of specific
71 interventions, the advent of new programming approaches implied the need to measure the
72 ‘implementation strength’ with which PHC systems deliver EBI packages and evaluate whether
73 improvement in implementation strength is associated with child survival (9–12).

74 The term ‘implementation strength’ has been used interchangeably with other terms, such as
75 ‘implementation intensity’ and ‘effective coverage’ and is defined here as a quantitative measure
76 of the ‘dose’, or the amount of input or activity, delivered to implement a program. Implementation
77 strength has a rich history in global health (13,14). Efforts to scale up emergency obstetric care led
78 to a ‘signal function framework’, which acknowledged that a set of performance indicators was
79 needed to evaluate maternal survival programs since the occurrence of obstetric emergencies was
80 rare within program settings (15). For similar reasons, researchers use signal functions to gauge
81 the effects of expanding access to abortion care services (16,17). The advent of ‘health systems
82 strengthening’ focused attention on the need to measure processes and outcomes of improving the

83 capacity of health systems to deliver PHC (18,19). The World Health Organization's 'health
84 systems building blocks' framework was central to this and provided a structure for monitoring
85 and evaluating health systems performance (20–22). This has included large-scale health facility
86 and household surveys in low- and middle-income countries, notably Service Provision
87 Assessments, Demographic and Health Surveys (DHS), Service Availability and Readiness
88 Assessments (SARA), and Multiple Indicator Cluster Surveys (23–25). In addition, researchers
89 have adapted common evaluation frameworks and applied them across health systems to measure
90 implementation strength (26,27). Approaches, such as the Balanced Score Card, provide a
91 dashboard of performance indicators that are used to detect variation and change in PHC
92 implementation strength (28–32). Similar approaches have been adapted to measure the same with
93 respect to domains of care, such as emergency obstetric and newborn care (EmONC) (33,34),
94 integrated management of childhood illness (IMCI) (12), and family planning (35).

95 *Implementation strength: Measurement and association with child health gains*

96 Efforts to apply the concept of implementation strength and evaluate its child health impacts have
97 faced methodological challenges. The first relates to derivation of the exposure variable. A
98 common approach has been to presuppose the intervention components that are most crucial to the
99 causal pathway and their relative importance based on intimate knowledge of the program,
100 literature review and credence to global measurement approaches (36,37). Researchers then design
101 methods and tools for capturing data on those components and perform simple or weighted
102 additive analysis to generate summary indices that represent implementation strength. Studies that
103 have used this approach to measure the individual-level child health response to implementation
104 strength have reported that there is a positive association between 'dose strength' of EBI
105 implementation and desirable MNCH behaviors (38,39). Although their exposure measurement

106 strategy is easy to conduct, theoretically justifiable and can produce results amenable to
107 comparison between geographies, it assumes that implementation strength is unidimensional, and
108 can create non-normal, skewed distributions. Thus, the additive summary score is not always
109 conceptually meaningful and may not accurately portray overall implementation strength (40,41).
110 Furthermore, this approach does not consider patterns between the wider, underlying set of
111 variables of program inputs and activities between units of analysis (e.g., facility, district). This is
112 problematic since these relationships contribute to the real variation between analytic units in terms
113 of the underlying construct that represents the exposure of interest, i.e., the implementation
114 strength that the health system exerted during the time and place of the evaluation at hand (42).
115 Neglecting this issue risks misallocating weight to given input and activity variables and omitting
116 potentially crucial ones from analysis (43). This, in turn, may undermine the quality of
117 associational analyses which assume that differences in measures of implementation explain
118 variation in health outcome.

119 Second, several analyses of the relationship between implementation strength and child health
120 have aggregated data on outcomes to the level of district, country or other clusters that represent
121 the level at which implementation actions are taken. Such studies have generated conflicting
122 results with respect to the association between dose strength of program implementation and child
123 health related care-seeking and child mortality (39,44). While this may be due to differences in the
124 strategy and content of programs, it's important to recognize that these studies are beset with
125 ecological biases (45). For example, one study in Uganda, which compared approaches for linking
126 individual-level service utilization indicators from household surveys with data on the provision
127 of maternal and newborn services at the facility- and ecological-level (district) to estimate levels
128 of effective coverage, found large discrepancies between estimates between the two approaches,

129 concluding that the ecological approach appreciably overestimates effective coverage (46). Third,
130 individual and ecological studies of implementation strength and child health have almost
131 exclusively depended on cross-sectional data (38,39,44). Although this often reflects the best use
132 of data available to address the research question, it precludes ascertainment of temporal order
133 between exposure to program intensity and the occurrence of the outcome. This temporal bias
134 raises questions about ‘reverse causality’ and prevents a ‘dose response’ interpretation of the
135 relationship between implementation strength and child health gains.

136 *Purpose and objectives*

137 The purpose of this paper is to demonstrate an analytic approach that overcomes these challenges.
138 We first develop summary indices of implementation strength using data from a SARA that was
139 conducted in three rural districts of Tanzania between May and August of 2011. SARA are
140 comprised of measures of general and domain specific attributes of PHC facilities’ structural
141 quality and are frequently implemented in low- and middle-income countries by special projects
142 or to assess health sector performance at national scale (47). Our analysis blends weighted-additive
143 methods and multivariate statistical procedures and applies them on a voluminous range of
144 indicators on MNCH provision. In doing so, we derive a minimal set of independent scales that
145 collectively represent variation in the intensity with which local health systems in the three districts
146 provided MNCH EBI. The analysis proceeds by linking scales of implementation strength with
147 longitudinal data that contain the survival trajectories of approximately 9,000 children, starting at
148 their birth from March to November 2011 in the catchment villages served by SARA facilities,
149 until the end of the cohort in December 2015. With this, we assess whether differences in the dose
150 of implementation strength exerted by PHC facilities on villages in which children experienced
151 infancy are associated with these children’s risk of dying during childhood.

152 In conducting this analysis, we address two objectives. The first is to validate the methodology we
153 use to measure implementation strength using routine facility data obtained through a popular data
154 collection methodology. We hypothesize that detection of an association between implementation
155 strength and child mortality, given that the SARA measured the provision of child survival EBI
156 that have been proven repeatedly and in various contexts, indicates that our measurement approach
157 is worthy of consideration for uptake by other studies that use SARA or similar data to evaluate
158 health system performance. Our second objective is to identify the components of the Tanzanian
159 integrated PHC program that drove child mortality reduction and give insight on inputs and
160 activities to emphasize to maximize program impact on child survival.

161 **Materials and Methods**

162 *The Tanzanian health system*

163 Since its founding, Tanzania has demonstrated commitment to achieving universal access to PHC.
164 In the 1970s, the country launched a national expansion of PHC facilities, called dispensaries and
165 rolled out a national village health worker program (48). During this period, the Ministry of Health
166 promulgated a policy that guaranteed mothers and children free access to basic services such as
167 immunization, nutrition, antenatal and postnatal care, growth monitoring, and treatment for minor
168 illness (49). In 2007, the Ministry of Health initiated the Primary Health Care Services
169 Improvement Program, which accelerated the expansion of dispensaries and higher-level PHC
170 facilities, called health centers, across the country (50). Since then, strategies to guide
171 implementation of this policy have emphasized making the above essential service package, as
172 well as skilled obstetric care, the IMCI, and family planning available at the health center and
173 dispensary level (51–55). The implementation of these strategies co-occurred with reduction in

174 child mortality in Tanzania, which between 2000 and 2015 declined from 130 to 58 per 1,000 live
175 births (56).

176 *Study environment and data sources*

177 Data for this study were obtained from the *Connect Project*. *Connect* was a cluster-randomized
178 controlled trial of the impact of community health workers on child survival that was conducted
179 by the Ifakara Health Institute in Tanzania from 2011- 2015 (57,58). *Connect* was situated in the
180 sentinel areas of the Ifakara and Rufiji Health and Demographic Surveillance Systems (HDSS) in
181 three rural districts (59,60). The Ifakara HDSS is in Morogoro, a landlocked region in southwestern
182 Tanzania, and is approximately 500 km from Dar es Salaam. The Rufiji HDSS is in Rufiji district
183 on the Indian Ocean coast approximately 150 km south of Dar es Salaam. The villages within these
184 sites are largely agrarian and depend on subsistence agriculture, fishing, and petty trading. See
185 Figure 1.

186 **INSERT FIGURE 1**

187 **Fig 1. The Connect Project Study Areas**

188 Between April and September 2011 *Connect* conducted a household survey in 101 villages to
189 obtain baseline measures of key MNCH behavior and service utilization indicators. At this time,
190 *Connect* also carried out a health facility survey using the SARA tools in the 109 health facilities
191 in the three districts to understand the context of MNCH service provision. Participants were
192 eligible for enrollment in the household survey if they were resident in a household randomized to
193 either study arm of the *Connect* trial, female and between the age of 18 and 49 or the caretaker of
194 an under-five year-old child. Potential participants were identified using the census data from the
195 HDSS and randomly selected within each village. Data collectors approached them in their homes

196 to recruit them for the study. Health facilities were eligible for the SARA if district health
197 authorities reported that they provided MNCH services and were in the study districts (Fig. 1).
198 Data collectors and district health authorities visited each of these facilities to explain the study
199 and obtain facilities' staff cooperation with data collection. The SARA obtained information from
200 facilities' staff on the villages in their catchment areas. Data collectors used geographic positioning
201 system tracking devices and mapped facility-to-village travel distances in a geographic
202 information system database (61,62).

203 *Connect* leveraged the longitudinal HDSS platforms for its impact evaluation. Individuals were
204 eligible for participation in HDSS data collection if they were residents of a household in any of
205 the 101 villages under demographic surveillance, not a minor, and could report on the occurrence
206 of key demographic events (births, deaths, in- and out-migrations) that took place in the household.
207 Until December 2015, staff of both HDSS visited households in the sentinel areas every four
208 months to collect information on household members, their relationships, ages, and sexes; births,
209 deaths, and in- and out-migration of household members. Every 1-2 years, censuses were
210 undertaken to enumerate old and new households. The censuses obtained data on maternal
211 education attainment and household income and assets (63,64). For our analysis, we linked
212 household survey, SARA and HDSS data by village name.

213 *Creating 'effective coverage scales' for specific domains of MNCH*

214 We followed a multi-step process to develop scales of MNCH implementation strength that was
215 exerted by PHC facilities in the *Connect* study areas on the villages that they served. For this, we
216 used data from four modules of the SARA (excluding the fifth module which obtained data on
217 implementation costs) and the household survey. We restricted our sample to the PHC facilities

218 (dispensaries and health centers, not hospitals) to which residents from the 101 study villages
219 would go for care.

220 The four SARA modules compiled categorical indicators on multiple domains of care: (1) general
221 facility readiness (staffing levels, management practices and infrastructure), (2) family planning,
222 (3) antenatal care, (4) intrapartum care, (5) postnatal care, (6) preventive services for children (e.g.,
223 immunizations, insecticide treated nets, counseling, assessment, classification components of
224 IMCI) and (7) sick childcare (trained staff and supplies to care for respiratory illness, diarrheal
225 disease, malaria, malnutrition). For indicators with more than one response category, we created
226 dummy variables so that all indicators used in the analysis were binary. For all indicators in each
227 domain, we calculated Cronbach's alpha coefficients to establish internal consistency and found
228 that all sets of indicators reported coefficients of 0.8 or higher (65). For each domain, we calculated
229 'effective coverage indices' using a weighted average approach. For each facility, we created the
230 general facility readiness effective coverage index by summing the values for each of the indicators
231 on staffing levels, management practices and infrastructure availability, respectively, and dividing
232 each total by the number of indicators related to each of these sub-categories. We summed these
233 averages and divided this sum by three, the number of categories in this domain.

234 We followed a similar procedure for the six service specific domains. Within each domain for each
235 facility, we summed the values of indicators that fell within three common sub-categories: (i) staff
236 training on domain-specific skills, (ii) the range and frequency with which domain specific
237 services were available at the facility, and (iii) current availability and recent stock outs of domain-
238 specific supplies, medicines, and equipment. We divided each of these sums by the total number
239 of indicators related to each sub-category, then summed those averages and divided this sum by
240 three. To incorporate coverage into this score, we identified the villages that were in each facility's

241 catchment area , and used data from the household survey to calculate village-level averages of
242 met need for family planning, and, with respect to respondents' most recent birth, receipt of ≥ 4
243 antenatal care visits, facility-based delivery, postnatal care, immunizations and, for respondents
244 with children that had recently had diarrheal, respiratory or febrile illness, receipt of needed
245 medications. We aggregated these averages to estimate the 'catchment specific coverage' for each
246 domain. We then multiplied the six measures of 'catchment specific coverage' by their
247 corresponding domain-specific scores for each facility. With this, for each facility, we obtained
248 seven domain specific effective coverage indices.

249 *Combining domain specific effective coverage indices to reduce data into independent,*
250 *parsimonious scales of implementation strength*

251 To obtain a smaller set of scales of implementation strength, we used principal components
252 analysis (PCA). We chose PCA because of its ability to reduce the highly correlated effective
253 coverage indices into fewer orthogonal principal components (PC), or scales, that maximize the
254 variation in the data and represent facilities' relative position in terms of implementation strength
255 (66). We determined the number of PC to retain in our analysis via parallel analysis. Per
256 convention, we retained PC with an eigenvalue of greater than or equal to 1 (42). To interpret each
257 PC, we examined the factor loadings and cosine² values that were reported for each of the
258 indicators used to formulate them. The higher loadings and cosine² values indicated greater and
259 higher-quality representation of each domain specific effective coverage index to each PC (67,68).

260 Next, we created scales that represented the levels of implementation strength to which each study
261 community was exposed. For this, we grouped facilities together if SARA respondents reported
262 that members of the same villages went to them for care or if the facility was within five kilometers
263 of the same villages. We combined the PC values of facilities in the same group using three-level

264 Bayesian mixed effects models with villages nested within catchment areas, and catchment areas
265 nested within districts, and fixed effects to denote villages distance from the nearest facility and
266 population size (69). Functionally, these models used the PC values of each facility for each
267 retained scale as the prior distribution to produce a posterior distribution of values that represented
268 the overall levels of implementation strength that facility groups exerted collectively (70). The
269 benefit of this approach is that it borrows information from facilities within each group to estimate
270 mean implementation strength scores that are shrunk to a central value, which results in more
271 stable estimates with smaller standard errors (71).

272 *Estimating the association between three dimensions of implementation strength on newborn and*
273 *child mortality*

274 We merged the combined scale values that had been assigned to villages to a longitudinal dataset
275 from the HDSS that included the survival trajectories and individual- and household level
276 covariates for children that were born between March and November 2011. We then estimated the
277 relationship between the implementation strength scores exerted by facility groups and the risk of
278 child mortality, modelling the implementation strength scales in their continuous form. To address
279 potential confounding, we incorporated fixed effects for covariates at the child level (child sex,
280 birth order, previous and subsequent birth interval durations), mother-level (age, marital status,
281 years of schooling), household socio-economic status (SES) ranking (1-5), and contextual
282 variables (distance from village to nearest health facility, whether a community health worker
283 worked in the village, HDSS zone in which villages were nested). To select our modelling strategy,
284 we conducted the Schoenfeld test of residuals to determine whether the assumption of proportional
285 hazards was met (72). Because multiple covariates marginally failed to satisfy this requirement,
286 we elected to use Weibull parametric hazard regression models, which capture the underlying

287 hazard of child mortality that is known to be high during the newborn period and decline as
288 children age. Children’s survival durations took into account the possibility of exit from the cohort
289 due to outmigration or death, and observations that experienced neither outcome by December 31,
290 2015 were subject to right censoring during analysis. To account for clustering of births within
291 villages, we incorporated in our models a random effect for village.

292 In our longitudinal dataset, there was missing data for two covariates, household SES and mothers’
293 years of schooling, for roughly one-fourth of the sample. After examination, we determined that
294 the omission of values reflected a ‘missingness at random’ pattern. To address this challenge, we
295 compared three approaches: First, we omitted the two covariates from our models and conducted
296 a complete case analysis; Second, we imputed missing values with the community-level median
297 value for the two covariates; Third, we used multiple imputation with chain equations (MICE)
298 (73). The results that these approaches produced were virtually identical as shown in Supplemental
299 File 1 (S1 Appendix). Therefore, we report the results of the third approach. We conducted the
300 entire analysis in *R Studio 4.0.2*.

301 Approval for this study was granted by the ethical review boards of the Ifakara Health Institute
302 (IHI/IRB/No. 16-2010), the National Institute for Medical Research of Tanzania
303 (NIMR/HQ/R.8a/Vol.IX/1203), and the Institutional Review Board of Columbia University
304 Medical Center (Protocol AAF3452). To obtain the data that was used for this study, the study
305 team first obtained information consent from human participants. Regarding the SARA and
306 household survey, which took place from April to September 2011, to enroll subjects, data
307 collectors trained and employed by *Connect* read an informed consent form aloud to potential
308 participants. Those that agreed to participate signed the form or provided an inked thumbprint. A
309 data collection team member that worked for the respective District Health Management Teams

310 witnessed each enrollment episode and provided a countersignature. Data collection on the
311 longitudinal cohort was managed by the HDSS teams of the Ifakara Health Institute in Kilombero
312 and Rufiji, respectively. To ensure compliance with their mandated data collection procedures, at
313 each follow up visit each HDSS respondent was provided with an explanation of the research
314 purpose, interviewing process and potential uses of participants' demographic and household
315 information. Those that agreed to participate, then provided written informed consent under the
316 observation of HDSS staff. This study did not enroll minors. Regarding the SARA and household
317 survey, information that could identify individual participants with the data they provided was kept
318 on the informed consent forms. These forms were transferred to a secure and private location at
319 the Ifakara Health Institute as soon as possible after each episode of data collection. After forms
320 were transferred to this location, the authors had no access to information that could identify
321 individual participants. Datasets availed to the authors from the HDSS for this analysis were
322 completely deidentified at the time of data transfer.

323 **Results**

324 Based on information from SARA respondents and travel distances, we found that the 101 study
325 villages sought PHC from 56 primary health care facilities. See Figure 2 and Table 1.

326 **INSERT FIGURE 2**

327 **Fig 2: Location of health care facilities in the study area**

328 In total, 234 indicators from the SARA were used to create domain specific effective coverage
329 indices (47 indicators on general facility readiness, 23 family planning services, 22 antenatal care,
330 54 delivery services, 23 postnatal care, 31 on preventive childhood services, 34 on sick child care).
331 Table 1 presents descriptive data on the domain specific effective coverage indices calculated for
332 the facilities in our sample. Of the 56 facilities, 6 were health centers, and 50 were dispensaries.

333 Thirty-five of the facilities were in the sentinel areas of the Ifakara HDSS (12 in Ifakara Rural, 5
 334 Ifakara Urban and 18 Ifakara Expansion) and 21 in the Rufiji HDSS (17 in Rufiji Rural, 4 Rufiji
 335 Rural). Supplemental File 2 contains a table that presents the median and range of scores of the
 336 effective coverage indices by HDSS zone (S2 Appendix).

337 **Table 1: Background characteristics of primary health care facilities (n = 56)**

	n (%)
Number of primary health care facilities retained	56 (100)
Facility type	
Health Center	6 (11)
Dispensary	50 (89)
Managing authority of health facility	
Government/public	36 (65)
Faith-based	14(25)
Military	3 (5)
Private	3 (5)
Geographic zone	
Ifakara Rural	12(21)
Ifakara Urban	5(9)
Ifakara Expansion ¹	18 (32)
Rufiji Rural	17 (30)
Rufiji Urban	4 (7)

338 Supplemental File 3 shows the Scree plot that was produced by the PCA and guided our parallel
 339 analysis (S3 Appendix). The first three PC reported eigenvalues of 1 or greater and together
 340 accounted for 77% of the variance among the seven domain specific effective coverage indices,
 341 whereas the remaining four scales held appreciably less explanatory potential. Thus, we chose to
 342 retain the first three scales only.

344 Figures 3a and 3b reflect the factor loadings and cosine² values for each PC, which reflect the
 345 relative contribution of each underlying effective coverage index to the retained scales. We found
 346 that the first PC, which accounted for 46% of variation in the effective coverage indices, was an

¹ Expansion refers to the geographic area comprising of communities in which the Ifakara HDSS was expanded in 2010.

347 implementation strength scale that reflected facilities' relative position in terms of the availability
348 and readiness to provide preventive health services, mostly child-related, with high loadings and
349 cosine² values for general facility readiness, antenatal care, postnatal care, and preventive
350 childhood services. The second PC, which accounted for 18% of the variation, represented a scale
351 of facilities' readiness to provide sick childcare and, to a lesser extent, family planning services;
352 and the third PC, which accounted for 13% of the variation, represented a scale of facilities'
353 readiness to provide intrapartum care.

354 **INSERT FIGURES 3A AND 3B**

355 **Fig 3a and 3b: Correlation plot of factor loadings (3a) and Cosine² values (3b) produced by PCA of MNCH**
356 **effective coverage indices.**

357 We used information provided by SARA respondents and geographic data to determine which of
358 the 56 facilities in our sample served the 101 villages in the study area respectively. In doing so,
359 we identified 89 different combinations of facilities. On average, villages were within the
360 catchment areas of three primary care facilities (range: 1, 7). For each village, we computed the
361 collective implementation strength that was exerted on them by the group of facilities in whose
362 catchments they were located using a Bayesian mixed effects model and estimating the empirical
363 Bayes mean implementation strength scores for the three scales that we retained. Figures 4a-c
364 illustrates the distribution of implementation strength scores assigned to the 89 different facility
365 groups that defined the catchment areas in which the 101 study communities were nested.

366 **INSERT FIGURES 4A-C**

367 **Fig 4a-c: Distribution of implementation scores across primary health care facility catchment areas in the study**
368 **area (n = 89)**

369

370 Supplemental Files 4a-c illustrates the geographic variation in implementation strength, comparing
 371 the distribution of implementation strength scores between groups of villages that were in five
 372 administrative zones within the study area, three in the Ifakara HDSS (Ifakara rural, urban and
 373 expansion areas) and two in Rufiji (rural and urban) (S4 Appendix). We observe the greatest
 374 variation in median implementation strength scores among zones for the first scale, which
 375 represents the availability and readiness to provide preventive care services for newborns and
 376 children as well as antenatal care, and the least variation between zones for the second scale, which
 377 represents access to sick childcare and family planning services.

378 Table 2 describes the characteristics of 8,999 children that were born from March-November 2011,
 379 the period spanning approximately three months before and after the start and end of the SARA,
 380 in the villages whose residents report to the above-described 56 facilities for care. Among these
 381 children, 526 (5.8%) died before December 31, 2015, when the follow up of the cohort ended. The
 382 median period of observation of children in the cohort was 46 months (range: 0, 58).

383 **Table 2: Characteristics of children born in catchment areas of the 56 primary health care facilities in the**
 384 **study area between March and November 2011 (n = 8,999)**

Characteristic	n (%)	Under-five year-old deaths (n, %)
Sex		
Male	4541 (50.1)	293 (55.7)
Female	4458 (50.0)	233 (44.3)
Birth order		
1	6193 (68.8)	365 (69.4)
2	1420 (15.7)	91 (17.3)
3	757 (8.4)	39 (7.4)
≥4	629 (7.0)	31 (5.9)
Previous birth interval duration		
First born child	6284 (69.8)	387 (73.4)
≤18 months	150 (1.7)	12 (2.3)
19-24 month	292 (3.2)	17 (3.3)
25-36 months	721 (8.0)	35 (6.7)
≥36 months	1551 (17.2)	75 (14.3)
Subsequent birth interval duration		
Last born child	7048 (78.3)	315 (59.9)

≤18 months	164 (1.8)	67 (12.7)
19-24 month	255 (2.8)	36 (6.8)
25-36 months	778 (8.8)	61 (11.6)
≥36 months	754 (8.4)	47 (8.9)
Mother's age at birth		
<20	1430 (15.9)	84 (16.0)
20-24	2173 (24.1)	121 (23.0)
25-29	2193 (24.4)	110 (20.9)
30-34	1679 (18.7)	113 (21.5)
35-39	1044 (11.6)	71 (13.5)
≥40	480 (5.3)	27 (5.1)
Mother marital status at birth		
Married/in union	5900 (65.6)	313 (59.5)
Single	3099 (34.4)	213 (40.5)
Mother number years of schooling		
None	1824 (20.3)	134 (25.5)
1-6 years (primary)	1064 (11.8)	66 (12.5)
≥ 6 (some secondary or more)	3283 (36.5)	197 (37.5)
Unknown ²	2798 (31.3)	129 (24.5)
Household SES at birth (quintile ranking)		
First (lowest)	1313 (14.6)	80 (15.2)
Second	1202 (13.4)	59 (11.2)
Third	1368 (15.2)	80 (15.2)
Fourth	1284 (14.3)	62 (11.8)
Fifth	1199 (13.3)	74 (14.1)
Unknown ³	2633 (29.2)	171 (32.3)
Distance to nearest health facility		
<5 kilometers	4839 (53.8)	271 (51.5)
5-10 kilometers	2719 (30.2)	164 (32.2)
>10 kilometers	1441 (16.0)	91 (17.3)
HDSS zone		
Ifakara Rural	3490 (38.8)	185 (35.2)
Ifakara Urban	939 (10.4)	71 (13.5)
Ifakara Expansion	2346 (26.1)	160 (30.4)
Rufiji Rural	1185 (13.2)	62 (11.8)
Rufiji Urban	1039 (11.5)	48 (9.1)

385 ² 2798 were missing, ³ 2633 were missing.

386

387 Table 3 presents the results of our analysis of the relationship between the three scales which

388 represent the implementation strength of MNCH to which children in the study area were exposed

389 during early infancy in 2011 and the risk of their dying before December 31, 2015. The first model

390 is an unadjusted analysis of the effect of the three dimensions of implementation strength on child
 391 mortality, the second the same analysis however adjusting for individual mother- and child-level
 392 covariates, and the third the same analysis however adjusting for individual mother- and child-
 393 level and contextual covariates. From our analysis, we observe an association between the first
 394 scale, which gauges the strength of general facility readiness and antenatal, postnatal, and
 395 preventive early childhood health care services accessible to the population, and the risk of dying
 396 during childhood. After adjustment for potentially confounding factors, the third model reports
 397 that for each unit increase in the implementation strength of these interventions, the risk of dying
 398 during childhood is 0.62 times lower (HR: 0.62; 95% CI: 0.40, 0.97).

399 **Table 3: Associations between child mortality risk and implementation strength scores reported from Model**
 400 **1 (adjusted, scores 1-3 only), Model 2 (scores and mother- and child-level covariates), and Model 3 (scores**
 401 **and mother- and child-level covariates and contextual-level covariates) (n = 8,999)**

	Model 1		Model 2		Model 3	
	HR	95% CI	HR	95% CI	HR	95% CI
Implementation strength scores						
Score 1 (ANC, PNC, and preventive child care)	0.76	0.60, 1.00	0.70*	0.54, 0.90	0.62*	0.40, 0.97
Score 2 (sick child care)	0.74	0.43, 1.25	0.60*	0.35, 0.99	0.56*	0.31, 0.99
Score 3 (intrapartum care)	0.86	0.43, 1.73	0.62	0.31, 1.23	0.47	0.19, 1.22
Child sex						
Female	-	-	-	-	-	-
Male	-	-	1.25*	1.05, 1.48	1.25*	1.05, 1.49
Birth order						
No. children (cont.)	-	-	0.99	0.88, 1.13	1.00	0.88, 1.13
Previous birth interval						
Months (cont.)	-	-	1.00	0.99, 1.00	1.00	0.99, 1.01
Subsequent birth interval						
Months (cont.)	-	-	1.01***	1.00, 1.02	1.01***	1.00, 1.02
Mother age at birth						
Years (cont.)	-	-	1.01	0.99, 1.02	1.01	1.00, 1.02
Mother marital status at birth						
Married/in union	-	-	-	-	-	-
Single	-	-	1.44**	1.20, 1.72	1.43**	1.19, 1.72
Mother number years of schooling						
Year of schooling (cont.)	-	-	0.97 [□]	0.94, 1.00	0.97	0.95, 1.01
Household SES at birth (quintile ranking)						
Fifth	-	-	-	-	-	-

Fourth	-	-	0.78 [□]	0.60, 1.02	0.79 [□]	0.60, 1.03
Third	-	-	0.96	0.74, 1.26	0.99	0.75, 1.29
Second	-	-	0.92	0.70, 1.22	0.95	0.71, 1.26
First	-	-	0.91	0.68, 1.21	0.93	0.70, 1.25
Distance to nearest health facility						
Kilometers (cont.)	-	-	-	-	1.14 [□]	0.98, 1.32
HDSS zone						
Ifakara Expansion	-	-	-	-	-	-
Ifakara Rural	-	-	-	-	0.85	0.58, 1.22
Ifakara Urban	-	-	-	-	1.25	0.81, 1.95
Rufiji Rural	-	-	-	-	1.12	0.70, 1.81
Rufiji Urban	-	-	-	-	0.95	0.54, 1.68
Community Health Worker Deployed to Community						
Yes					0.93	0.76, 1.13

402 □ = p-values < 0.1, * = p-values < 0.05, ** = p-values < 0.01, *** = p-values < 0.001

403 The third model suggests a more protective effect of implementation strength on child mortality
404 with regards to the second scale, which represents the availability and readiness of nearby facilities
405 to provide sick child care. After adjustment for multi-level covariates, we noted that among
406 children in our cohort, increases exposure to higher levels of sick child care implementation
407 intensity were associated with, on average, with 0.56 times lower risk of dying during childhood
408 (HR: 0.56, 95% CI: 0.31, 0.99). There was no statistically significant association detected between
409 variation in the third scale of implementation strength, which represents the strength of intrapartum
410 care services available to the population, and child mortality risk. In addition to evaluating the
411 child mortality response, we adapted Models 1-3 so that they could indicate whether variation in
412 implementation strength, as represented by the three scales, were associated with children's risk
413 of dying in the first month and first year of life, respectively. These analyses reported hazards
414 ratios of magnitudes that were similar to those reported by Model 3, but these results were not
415 statistically significant (S5 Appendix).

416 Discussion

417 Our analysis reports that children exposed to higher levels of implementation strength of
418 preventive and curative childcare, including antenatal care, postnatal care, and family planning
419 services available to mothers, from the PHC facilities that served their community at the time of
420 birth and early infancy were less likely to die during childhood than those exposed to lower levels
421 of implementation strength. Whereas the same analyses of the relative risk of newborn and infant
422 mortality associated with implementation strength did not reveal a significant relationship, we
423 believe that this owes the limitations in the power accorded by the sample size and number of
424 mortality events in the first month and year of life in our cohort. Furthermore, we observe no
425 association between child mortality and the third implementation strength scale, which represents
426 the strength of intrapartum care services. This finding contradicts similar studies which measured
427 the health response to obstetric care quality. For example, Tiruneh et al. (2018). in Ethiopia used
428 an additive approach to develop a gradient of EmONC implementation strength and conducted an
429 analysis which reported a positive association between higher levels of EmONC implementation
430 strength and facility-based delivery and met need for EmONC (34). Although our finding may be
431 surprising, it is important to recall that the third scale explained a relatively small proportion of
432 the overall variance among the effective coverage indices that we reduced into our independent
433 gradients of implementation strength. If, in our analysis of SARA data, we had included hospitals
434 and higher-level health centers, which are relatively better equipped to handle labor and delivery,
435 it is likely that the sample would have contained more variation with respect to this domain of
436 care. In turn, this might have led to a set of scales that demonstrated a more potent effect of
437 intrapartum care implementation strength.

438 Our findings contribute to the debate about the role of primary health care programs and services
439 in the precipitous child mortality decline that Tanzania experienced from 2000-2015. Importantly,

440 they corroborate the findings issued by Masanja et al. (2008) in their analysis of the trends and
441 drivers of child mortality using four consecutive DHS surveys between 1990 and 2008. This study
442 identified the synergistic effect of increases in public sector health expenditure, implementation of
443 decentralization policies and expansion in the coverage of high impact child health interventions,
444 such as insecticide treated nets (ITN), extended program immunizations (EPI), and IMCI, as
445 contextual factors important to the reduction of child mortality observed in Tanzania during this
446 period (74). Afnan-Holmes and coauthors, make similar observations when reviewing Tanzania's
447 achievement of Millennium Development Goal 4, also attributing success to increased coverage
448 of EPI, IMCI and ITN (75). Our analysis underscores the importance of investments in expanding
449 these interventions. Furthermore, our interpretation of the first set of factor loadings produced by
450 the PCA suggests that ensuring effective coverage of routine antenatal and postnatal care as well
451 as general facility readiness, in terms of regular staffing, general infrastructure and routine
452 management inputs and processes, may be as important to enhancing child survival prospects in
453 the population as preventive and promotive health interventions that target children under five.
454 This information is relevant to health system managers and planners in Tanzania as they prioritize
455 ways to invest limited resources in ensuring effective coverage of PHC services and scaling up.
456 This analysis illuminated an appreciable child mortality response to slight variation in the strength
457 of PHC performance within small geographic areas. In our examination we distinguished the
458 relative contributions of different domains of MNCH to this variability and provided insight on
459 how improvements vis-à-vis these domains can impact children's prospects of surviving
460 childhood. Yet, we did not explore the factors that explain why, within relatively small areas,
461 some facilities perform better than others. Previous studies in Tanzania have sought to understand
462 how differences between nearby facilities condition patient care seeking. For example, one study

463 reported that care seeking behaviors were shaped by differences between facilities in terms of the
464 quality of consultations and availability prescriptions, knowledge-level of staff, and availability of
465 physicians and essential supplies and equipment (76). Another study sought to understand the
466 drivers of variation between nearby facilities in rural areas in terms of provider competence and
467 practice quality, and reported that this was a function of facility ownership (private for profit, non-
468 governmental organization, public sector), population density of facilities' catchment area, and
469 health workers training, tenure and experience (77). Other studies in Tanzania on this topic have
470 focused on specific domains of MNCH. Kahabuka et al. (2011) attributed patient care seeking for
471 preventive and sick child care services to differences between nearby facilities in terms of the
472 availability of diagnostic equipment, essential medicines and skilled staff (78). Kanté et al. (2016)
473 reported that differences between facilities' performance of emergency obstetric care signal
474 functions conditioned women's care seeking for intrapartum care (79). Future research, in
475 Tanzania and similar settings, should continue to examine the determinants and processes that
476 generate within-small area variation in service readiness, availability and quality since this
477 information, as our findings suggest, could help address problems that underlie preventable child
478 deaths.

479 Effective coverage measures that link population level data on access to health care and health
480 outcomes with facility data on health care quality are increasingly reported (80–82). However,
481 there is little guidance on appropriate methods for linking, e.g., which data to link, for which units
482 and with what temporal alignment (46). Furthermore, decisions to link data sets are usually not
483 foreseen during study design, and, therefore, researchers must be opportunistic in the linkage
484 methods they employ. In our analysis, we linked individual-level data on child survival to
485 summary scores of MNCH coverage effectiveness of the facilities in the environs of children's

486 household and communities and found that the contextual effect of implementation strength in
487 these areas was associated in lower risks of child mortality. However, this might differ from an
488 analysis of the same relationship that linked individual outcomes with performance scores of exact
489 facilities where children obtained care. Future research is needed on effective coverage estimation
490 and its health effects that compares methods that operationalize coverage in geographic terms
491 versus in terms of where individuals utilized services. Identification of biases associated with
492 either linkage approach will help health systems researchers determine the data and data systems
493 requirements for measuring health systems strengthening and its impacts.

494 To our knowledge this is the only study that has benefitted from the ability to link data on ‘dose
495 delivered’ of implementation strength within local health systems with longitudinal data on the
496 survival of children nested in the underlying population. Although other studies, for example in
497 Ethiopia and Malawi, addressed similar questions related to implementation strength of MNCH
498 services, the larger geographic scope of their analysis and lack of prospective data on the outcome
499 compelled authors to draw upon repeat cross-sectional data from the DHS or special project
500 surveys, and perform ecological analyses, which are subject to biases (38,44). Whereas these
501 findings are valuable insofar as they evaluate the impact of the large-scale rollout of national child
502 survival programs, our analysis fills an important gap in that it reports the effects of variation in
503 dose delivered of implementation strength by routine delivery systems within a relatively small
504 area on individual level mortality risks that unfold over time after exposure. Future research should
505 seek opportunities to leverage existing longitudinal data collection platforms and integrate them
506 into investigations of the population-level health response to changes in health systems strength.

507 Our analysis is not without limitations. First, though we opted to use PCA as a data reduction
508 approach, we recognize that this approach has drawbacks. For example, there are examples that

509 have demonstrated that the use of PCA results in the misclassification of subjects vis-à-vis the
510 gradient of the underlying construct when the PCs used explained less than 30% of total variance
511 (83). Although our finding that higher levels of readiness and availability of sick child care was
512 associated with lower child mortality risk seems intuitive, it is not immune to this type of critique.
513 Second, despite our review of the risks of weighted-additive methods, we used that approach to
514 derive effective coverage indices of different domains of MNCH that we later subjected to PCA.
515 This came after reviewing alternative data reduction strategies, including use of PCA only to
516 reduce the 234 indicators in the SARA to implementation strength scales. In the end, we felt that
517 the weighted-additive approach, enabled us to represent the natural partition in our data between
518 domain-specific measures of MNCH availability and readiness and best identify which
519 components of MNCH care are most relevant to child mortality reduction. Third, there were
520 significant levels of missingness for two covariates in the HDSS, which supplied our data on child
521 survival. However, after comparing three approaches for addressing this problem (complete case
522 analysis, median imputation, and MICE) we found that comparable results were obtained under
523 all strategies, which indicated that missing data did not seriously affect the overall association
524 between implementation strength and child survival. Fourth, the analysis assumes that the levels
525 of implementation strength that we obtained from the cross-sectional SARA from May to
526 September 2011 reflect the longer period for which we subset the longitudinal data from the HDSS
527 (child survival trajectories starting at birth from March to November 2011). Fifth, our analysis
528 links children born during the period surrounding the SARA with data on the facilities in whose
529 catchment their communities were located; however, our analysis could not ascertain whether
530 these children ever sought health care at other health facilities during their first 4-5 years of life.
531 Finally, although the analysis was able to link within district variation in levels of implementation

532 strength exerted by primary health care facilities with longitudinal data on child survival, these
533 data are observational and, thus, fall short of permitting inference that is truly causal.

534 **Conclusion**

535 We developed gradients that quantify the strength with which MNCH services were delivered to
536 the populations in the local health systems in three districts of Tanzania in 2011, and, by linking
537 these scales with longitudinal survival trajectories of children born in the underlying population,
538 we evaluated whether variation in implementation strength at the time of childbirth and early
539 infancy were associated with mortality risk during childhood. The results suggest that the intensity
540 with which preventive care services, including general facility readiness, antenatal and postnatal
541 care, and preventive care for children, were made available to the population, as well as that of
542 curative care and family planning services, was associated with lower mortality risk. Since these
543 scales reflect the performance of interventions that are part of the Tanzania's essential MNCH
544 services package through routine delivery system channels, local health authorities can use these
545 metrics, which were derived by using data readily available from facilities, to better understand
546 the health impact of their implementation on the populations they serve. Additional research of
547 this nature should be undertaken using data from hospitals to obtain data with greater variation on
548 the performance of obstetric and gynecological services to better understand how health systems
549 meet the needs of women during the crucial intra- and postpartum periods. Health system managers
550 and decision-makers can use this information to inform planning, resource allocation and
551 implementation adjustments and maximize the impact of health system strengthening on
552 population health.

553

554

555 Supplemental information

- 556 1. Supplemental File 1 (S1 Appendix): Association between child mortality risk and MNCH
557 implementation strength using three approaches for addressing missing data for
558 covariates on household SES and mothers' years of schooling.
559 Model 1: Complete case analysis omitting covariates with missing data (n=6,068).
560 Model 2: Addresses missing data by imputing community-level median values of household SES and
561 mothers' years of schooling where data are missing (n=8,999).
562 Model 3: Addresses missing data by using multiple imputation with chained equations (n=8,999).
563
564
- 565 2. Supplementary File 2 (S2 Appendix): Effective coverage indices (median and range) for
566 primary health care facilities (n = 56).
567
568
- 569 3. Supplemental File 3 (S3 Appendix): Scree plot reported by the principal components
570 analysis of effective coverage indices.
571
572
- 573 4. Supplemental Files 4a (S4a Appendix): Comparison of the distribution of implementation
574 strength scale 1's values across geographic zones in the study area.
575
- 576 5. Supplemental Files 4b (S4b Appendix): Comparison of the distribution of
577 implementation strength scale 2's values across geographic zones in the study area.
578
- 579 6. Supplemental Files 4c (S4c Appendix): Comparison of the distribution of implementation
580 strength scale 3's values across geographic zones in the study area.
581
582
- 583 7. Supplemental File 5 (S5 Appendix): Associations between newborn (<1 month), infant
584 (<12 months) and child mortality (<60 months) risks and implementation strength scores
585 reported from Model 3 (IS scores, mother-, child- and contextual-level covariates)
586 (n=8,999)
587

588

589

590

591

592

593 **Acknowledgments**

594 We gratefully acknowledge the funding and advisory support of the Doris Duke Charitable
595 Foundation (grant number DDCF2009058a) and Comic Relief UK (grant number 112259), and
596 the contributions of staff of the Health and Demographic Surveillance Systems of the Ifakara
597 Health Institute including Amri Shamte, Matthew Alexander, Francis Levira, Eveline Geubbles,
598 Rose Nathan. We also recognize the contributions of Jitihada Baraka, Ruth Wilson, Awena
599 Malendo, Mustafa Ngozi and Ahmed Hingora. Finally, we recognize with gratitude the leadership
600 and importance of staff and supporters from Kilombero, Rufiji and Ulanga districts, including staff
601 of their respective Council Health Management Teams and authorities of the villages in which this
602 research was conducted.

603
604
605
606
607
608
609
610
611
612
613
614
615

616 **References**

- 617 1. Liu L, Oza S, Hogan D, Chu Y, Perin J, Zhu J, et al. Global, regional, and national causes
618 of under-5 mortality in 2000–15: an updated systematic analysis with implications for the
619 Sustainable Development Goals. *The Lancet*. 2016 Dec;388(10063):3027–35.
620 doi:10.1016/S0140-6736(16)31593-8
- 621 2. Taylor AW, Blau DM, Bassat Q, Onyango D, Kotloff KL, Arifeen S el, et al. Initial
622 findings from a novel population-based child mortality surveillance approach: a
623 descriptive study. *Lancet Glob Health*. 2020 Jul;8(7):e909–19. doi:10.1016/S2214-
624 109X(20)30205-9
- 625 3. Vakili R, Emami Moghadam Z, Khademi G, Vakili S, Saeidi M. Child Mortality at
626 Different World Regions: A Comparison Review. *Int J Pediatr*. 2015;3(4–2):809–16.
- 627 4. Hug L, Alexander M, You D, Alkema L. National, regional, and global levels and trends
628 in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a
629 systematic analysis. *Lancet Glob Health*. 2019;7(6):e710–20. doi: 10.1016/S2214-
630 109X(19)30163-9.
- 631 5. Bhutta ZA, Ali S, Cousens S, Ali TM, Haider BA, Rizvi A, et al. Interventions to address
632 maternal, newborn, and child survival: what difference can integrated primary health care
633 strategies make? *The Lancet*. 2008;372(9642):972–89. doi:10.1016/S0140-
634 6736(08)61407-5.
- 635 6. Kerber KJ, de Graft-Johnson JE, Bhutta ZA, Okong P, Starrs A, Lawn JE. Continuum of
636 care for maternal, newborn, and child health: from slogan to service delivery. *Lancet*.
637 2007; 370(9595):1358–69. doi:10.1016/S0140-6736(07)61578-5.
- 638 7. Dean S v, Lassi ZS, Imam AM, Bhutta ZA. Preconception care: closing the gap in the
639 continuum of care to accelerate improvements in maternal, newborn and child health.
640 *Reprod Health*. 2014; 11(3): 1-8. doi:10.1186/1742-4755-11-S3-S1.
- 641 8. Bryce J, el Arifeen S, Pariyo G, Lanata CF, Gwatkin D, Habicht JP. Reducing child
642 mortality: can public health deliver? *The Lancet*. 2003 Jul 12;362(9378):159–164.
643 doi:10.1016/S0140-6736(03)13870-6.
- 644 9. Heidkamp R, Hazel E, Nsona H, Mleme T, Jamali A, Bryce J. Measuring Implementation
645 Strength for Integrated Community Case Management in Malawi: Results from a National
646 Cell Phone Census. *Am J Trop Med Hyg*. 2015 Oct 7;93(4): 861. doi:10.4269/ajtmh.14-
647 0797.
- 648 10. Hazel E, Amouzou A, Park L, Banda B, Chimuna T, Guenther T, et al. Real-Time
649 Assessments of the Strength of Program Implementation for Community Case
650 Management of Childhood Illness: Validation of a Mobile Phone-Based Method in
651 Malawi. *Am J Trop Med Hyg*. 2015 Mar 4;92(3): 660. doi:10.4269/ajtmh.14-0396.

- 652 11. Miller NP, Amouzou A, Tafesse M, Hazel E, Legesse H, Degefie T, et al. Integrated
653 Community Case Management of Childhood Illness in Ethiopia: Implementation Strength
654 and Quality of Care. *Am J Trop Med Hyg.* 2014;91(2): 424: doi:10.4269/ajtmh.13-0751.
- 655 12. Bryce J, Amouzou A, Hazel E, Miller N, Johns B, Gilroy K, et al. Measuring the strength
656 of implementation of community case management of childhood illness within the
657 Catalytic Initiative to Save a Million Lives. Baltimore, MD; 2011 [cited 2021 Jun 6].
658 Available from: [https://www.jhsph.edu/departments/international-](https://www.jhsph.edu/departments/international-health/IIP/projects/catalytic_initiative/wp-implementation-strength.pdf)
659 [health/IIP/projects/catalytic_](https://www.jhsph.edu/departments/international-health/IIP/projects/catalytic_initiative/wp-implementation-strength.pdf)
[initiative/wp-implementation-strength.pdf](https://www.jhsph.edu/departments/international-health/IIP/projects/catalytic_initiative/wp-implementation-strength.pdf):
- 660 13. Hargreaves JR, Goodman C, Davey C, Willey BA, Avan BI, Armstrong JRM. Measuring
661 implementation strength : lessons from the evaluation of public health strategies in low-
662 and middle-income settings. *Health Policy and Planning.* 2016;(March):860–7.
- 663 14. Schellenberg J. Measuring implementation strength: Literature review draft report 2012.
664 London; 2012 [cited 2021 Jun 6]. Available from: [https://ideas.lshtm.ac.uk/wp-](https://ideas.lshtm.ac.uk/wp-content/uploads/2017/08/IDEAS-Measuring-implementation-strength-report.pdf)
665 [content/uploads/2017/08/IDEAS-Measuring-implementation-strength-report.pdf](https://ideas.lshtm.ac.uk/wp-content/uploads/2017/08/IDEAS-Measuring-implementation-strength-report.pdf)
- 666 15. Bailey P, Paxton A, Lobis S, Fry D. The availability of life-saving obstetric services in
667 developing countries: An in-depth look at the signal functions for emergency obstetric
668 care. *International Journal of Gynecology & Obstetrics.* 2006 Jun;93(3): 285-291.
669 doi:10.1016/j.ijgo.2006.01.028
- 670 16. Otsea K, Benson J, Alemayehu T, Pearson E, Healy J. International Journal of
671 Gynecology and Obstetrics Testing the Safe Abortion Care model in Ethiopia to monitor
672 service availability , use , and quality. *International Journal of Gynecology and Obstetrics.*
673 2011;115(3):316–21. doi.org/10.1016/j.ijgo.2011.09.003
- 674 17. Campbell OMR, Aquino EML, Vwalika B, Gabrysch S. Signal functions for measuring
675 the ability of health facilities to provide abortion services: an illustrative analysis using a
676 health facility census in Zambia. *BMC Pregnancy Childbirth.* 2016;16(1):1–13.
- 677 18. World Health Organization. World Health Report 2008: Primary Health Care, Now More
678 Than Ever. Geneva, Switzerland; 2008 [cited 2021 Jun 10]. Available from:
679 [https://reliefweb.int/report/world/world-health-report-2008-primary-health-care-now-](https://reliefweb.int/report/world/world-health-report-2008-primary-health-care-now-more-ever?)
680 [more-ever?](https://reliefweb.int/report/world/world-health-report-2008-primary-health-care-now-more-ever?).
- 681 19. Adam T, de Savigny D. Systems thinking for strengthening health systems in LMICs:
682 need for a paradigm shift. *Health Policy Plan.* 2012;27 Suppl 4:iv1-3.
683 doi:10.1093/heapol/czs084.
- 684 20. World Health Organization. Everybody’s business: strengthening health systems to
685 improve health outcomes: WHO’s framework for action. Geneva, Switzerland; 2007.
- 686 21. Boerma T, Abou-Zahr C, Bos E, Hansen P, Addai E, Low-Beer D. Monitoring and
687 evaluation of health systems strengthening. Geneva, Switzerland; 2010 [cited 2021 Jun 6].
688 Available from: https://www.who.int/healthinfo/HSS_MandE_framework_Nov_2009.pdf

- 689 22. World Health Organization. Monitoring the building blocks of health systems : a
690 handbook of indicators and their measurement strategies. Geneva, Switzerland; 2010
691 [cited 2021 Jun 13]. Available from:
692 <https://apps.who.int/iris/bitstream/handle/10665/258734/9789241564052-eng.pdf>.
- 693 23. ICF International Inc. Service Provision Assessment: Overview and Methodology. The
694 Demographic and Health Surveys Program. 2021 [cited 2021 Jun 13]. Available from:
695 <https://dhsprogram.com/methodology/Survey-Types/SPA.cfm>
- 696 24. Sheffel A, Karp C, Creanga AA. Use of Service Provision Assessments and Service
697 Availability and Readiness Assessments for monitoring quality of maternal and newborn
698 health services in low-income and middle-income countries. *BMJ Glob Health*.
699 2018;3(6):e001011.
- 700 25. Khan S, Hancioglu A. Multiple indicator cluster surveys: delivering robust data on
701 children and women across the globe. *Stud Fam Plann*. 2019;50(3):279–86.
- 702 26. Sherr K, Fernandes Q, Kanté AM, Bawah A, Condo J, Mutale W, et al. Measuring health
703 systems strength and its impact: Experiences from the African Health Initiative. *BMC*
704 *Health Serv Res*. 2017;17: 29-38. doi:10.1186/s12913-017-2658-5.
- 705 27. Bryce J, Requejo JH, Moulton LH, Ram M, Black RE. A common evaluation framework
706 for the African Health Initiative. *BMC Health Serv Res*. 2013;13(S2): 1-9.
707 doi:10.1186/1472-6963-13-S2-S10.
- 708 28. Kaplan R, Norton D. The balanced scorecard: translating strategy into action. Boston,
709 MA: Harvard Business School Press; 1996.
- 710 29. Niven P. Balanced scorecard step-by-step for government and nonprofit agencies.
711 Hoboken, NJ: John Wiley & Sons; 2003.
- 712 30. Inamdar N, Kaplan R, Bower M. Applying the balanced scorecard in healthcare provider
713 organizations. *Journal of Healthcare Management*. 2002;47:179–95.
- 714 31. Amer F, Hammoud S, Khatatbeh H, Lohner S, Boncz I, Endrei D. The deployment of
715 balanced scorecard in health care organizations: is it beneficial? A systematic review.
716 *BMC Health Serv Res*. 2022;22(1): 1-14. doi:10.1186/s12913-021-07452-7.
- 717 32. Bohm V, Lacaille D, Spencer N, Barber CE. Scoping review of balanced scorecards for
718 use in healthcare settings: development and implementation. *BMJ Open Qual*. 2021
719 Jul;10(3):e001293. doi:10.1136/bmjopen-2020-001293.
- 720 33. Zaidi S, Riaz A, Rabbani F, Azam SI, Imran SN, Pradhan NA, et al. Can contracted out
721 health facilities improve access, equity, and quality of maternal and newborn health
722 services? Evidence from Pakistan. *Health Res Policy Syst*. 2015; 13(S1): 47-56.
723 doi:10.1186/s12961-015-0041-8
- 724 34. Tiruneh GT, Karim AM, Avan BI, Zemichael NF, Wereta TG, Wickremasinghe D, et al.
725 The effect of implementation strength of basic emergency obstetric and newborn care

- 726 (BEmONC) on facility deliveries and the met need for BEmONC at the primary health
727 care level in Ethiopia. *BMC Pregnancy Childbirth*. 2018 Dec 2;18(1): 1-11.
728 doi:10.1186/s12884-018-1751-z
- 729 35. RamaRao S, Jain A. Constructing indicators for measurement and improvement of the
730 quality of family planning programs: An example using data on choice from the
731 Philippines, 1997-1998. In: Leisher S, Sprockett A, Longfield K, Montagu D, editors.
732 *Quality Measurement in Family Planning: Past, Present, Future: Papers from the Bellagio*
733 *Meeting on Family Planning Quality*. Oakland, CA: Metrics for Management; 2016. p.
734 47–61.
- 735 36. Wehrmeister FC, Barros AJD, Hosseinpoor AR, Boerma T, Victora CG. Measuring
736 universal health coverage in reproductive, maternal, newborn and child health: An update
737 of the composite coverage index. *PLoS One*. 2020 Apr 29;15(4):e0232350.
- 738 37. Barros AJD, Victora CG. Measuring Coverage in MNCH: Determining and Interpreting
739 Inequalities in Coverage of Maternal, Newborn, and Child Health Interventions. *PLoS*
740 *Med*. 2013 May 7;10(5):e1001390.
- 741 38. Karim AM, Admassu K, Schellenberg J, Alemu H, Getachew N, Ameha A, et al. Effect of
742 Ethiopia's Health Extension Program on Maternal and Newborn Health Care Practices in
743 101 Rural Districts: A Dose-Response Study. *PLoS One*. 2013;8(6).
744 doi:10.1371/journal.pone.0065160
- 745 39. Corsi DJ, Subramanian S v. Association between coverage of maternal and child health
746 interventions, and under-5 mortality: A repeated cross-sectional analysis of 35 sub-
747 Saharan African countries. *Glob Health Action*. 2014;7(1).
- 748 40. Brown L, Tyane M, Bertrand J, Lauro D, Abou-Ouakil M, Demaria L. Quality of care in
749 family planning services in Morocco. *Stud Fam Plann*. 1995;154–68.
- 750 41. Bellows B, Behl R, Abuya T, Muriuki A, Bajracharya A, Choi Y. Benchmarking to assess
751 quality of family planning services: Construction and use of indices for family planning
752 readiness in Kenya with data from 2010 and 2014. In: Leisher SH SA, Longfield K,
753 Montagu D, editors. *Quality Measurement in Family Planning: Past, Present and Future:*
754 *Papers from the Bellagio Meeting on Family Planning Quality*. Oakland, CA: Metrics for
755 Measurement; 2015.
- 756 42. Abdi H, Williams LJ. Principal component analysis. *Wiley Interdiscip Rev Comput Stat*.
757 2010 Jul;2(4): 433-459. doi:10.1002/wics.101.
- 758 43. Mallick L, Temsah G, Wang W. Comparing summary measures of quality of care for
759 family planning in Haiti, Malawi, and Tanzania. *PLoS One*. 2019 Jun 7;14(6):e0217547.
- 760 44. Amouzou A, Kanyuka M, Hazel E, Heidkamp R, Marsh A, Mleme T, et al. Independent
761 Evaluation of the integrated Community Case Management of Childhood Illness Strategy
762 in Malawi Using a National Evaluation Platform Design. *Am J Trop Med Hyg*. 2016 Mar
763 2;94(3): 574. doi:10.4269/ajtmh.15-0584.

- 764 45. Greenland S, Morgenstern H. Ecological Bias, Confounding, and Effect Modification. *Int*
765 *J Epidemiol*. 1989;18(1):269–74.
- 766 46. Willey B, Waiswa P, Kajjo D, Munos M, Akuze J, Allen E, et al. Linking data sources for
767 measurement of effective coverage in maternal and newborn health: what do we learn
768 from individual-vs ecological-linking methods? *J Glob Health*. 2018;8(1).
- 769 47. The SARA Project Management Group. Service Availability and Readiness Assessment
770 (SARA), An annual monitoring system for service delivery: Reference Manual. Geneva;
771 2013.
- 772 48. Heggenhougen K, Vaughan P, Muhondwa E, Rutabanzibwa, Ngaiza J. Community health
773 workers: the Tanzanian experience. *Community health workers: the Tanzanian*
774 *experience*. New York, Oxford University Press; 1987.
- 775 49. Jonsson U. Ideological framework and health development in Tanzania 1961–2000. *Soc*
776 *Sci Med*. 1986 Jan 1;22(7):745–53.
- 777 50. The United Republic of Tanzania Ministry of Health and Social Welfare. The Primary
778 Health Care Improvement Programme (MMAM). Dar-es-Salaam; 2007.
- 779 51. The United Republic of Tanzania Ministry of Health and Social Welfare. The National
780 Road Map Strategic Plan to Accelerate Reduction of Maternal, Newborn and Child Deaths
781 in Tanzania (2008-15). The National Road Map Strategic Plan to Accelerate Reduction of
782 Maternal, Newborn and Child Deaths in Tanzania (2008-15). Dar es Salaam, Tanzania;
783 2008.
- 784 52. The United Republic of Tanzania Ministry of Health and Social Welfare J 2009 J. Health
785 Sector Strategic Plan III. Dar es Salaam, Tanzania; 2009 May.
- 786 53. Masanja H, Schellenberg JA, de Savigny D, Mshinda H, Victora CG. Impact of Integrated
787 Management of Childhood Illness on inequalities in child health in rural Tanzania. *Health*
788 *Policy Plan [Internet]*. 2005; 20(Suppl 1): i77–84.
- 789 54. Tanzania IMCI Multi-Country Evaluation Health Facility Survey Study Group. The effect
790 of Integrated Management of Childhood Illness on observed quality of care of under-fives
791 in rural Tanzania. *Health Policy Plan*. 2004; 19(1): 1-10.
- 792 55. Bryce J, Gouws E, Adam T, Black RE, Schellenberg JA, Manzi F, et al. Improving quality
793 and efficiency of facility-based child health care through Integrated Management of
794 Childhood Illness in Tanzania. *Health Policy Plan*. 2005; 20(Suppl 1): i69–76.
- 795 56. The World Bank. Under-five Mortality Trends - Tanzania [Internet]. [cited 2021 Nov 3].
796 Available from:
797 https://data.worldbank.org/indicator/SH.DYN.MORT?locations=TZ&year_high_desc=true
798 e
- 799 57. Ramsey K, Hingora A, Kante M, Jackson E, Exavery A, Pemba S, et al. The Tanzania
800 Connect Project: a cluster-randomized trial of the child survival impact of adding paid

- 801 community health workers to an existing facility-focused health system. *BMC Health Serv*
802 *Res.* 2013 May;13(2): 1-14.
- 803 58. Kanté AM, Exavery A, Jackson EF, Kassimu T, Baynes CD, Hingora A, et al. The impact
804 of paid community health worker deployment on child survival: The connect randomized
805 cluster trial in rural Tanzania. *BMC Health Serv Res.* 2019;19(1): 1-14.
- 806 59. Geubbels E, Amri S, Levira F, Schellenberg J, Masanja H, Nathan R. Health and
807 Demographic Surveillance System Profile: The Ifakara Rural and Urban Health and
808 Demographic Surveillance System (Ifakara HDSS). *Int J Epidemiol.* 2015 Jun;44(3):848–
809 61.
- 810 60. Mrema S, Kante AM, Levira F, Mono A, Irema K, de Savigny D, et al. Health and
811 Demographic Surveillance System Profile: The Rufiji Health and Demographic
812 Surveillance System (Rufiji HDSS). *Int J Epidemiol.* 2015 Apr;44(2):472–83.
- 813 61. Noronha JS, Kante AM, Hingora AI, Phillips JF. Geographical disparities of cause-
814 specific under-5 mortality in Rufiji District, Coastal Region, Tanzania: a longitudinal-
815 based study. *The Lancet.* 2013 Jun 17;381:S99.
- 816 62. Kante A, HELLERINGER S, Mrema S, Levira F, Masanja H, Nathan R, et al. Does proximity
817 to health facilities improve child survival? New evidence from a longitudinal study in
818 rural Tanzania. *Dar es Salaam*; 2013.
- 819 63. Mwangeni E, Masanja H, Juma Z, Momburi D, Mkilindi Y, Mbuya C, et al. Socio-
820 economic Status and health Inequalities in Rural Tanzania: Evidence from the Rufiji
821 Demographic Surveillance System. In: INDEPTH Network, editor. *Measuring Health
822 Equity in Small Areas.* Aldershot, England: Ashgate Publishing, Ltd.; 2006. p. 19–33.
- 823 64. Nathan R, Armstrong-Schellenberg J, Masanja H, Charles S, Mukasa O, Mshinda H.
824 Child Health Inequity in Rural Tanzania: Can the National Millennium Development
825 Goals Include the Poorest? In: INDEPTH Network, editor. *Measuring Health Equity in
826 Small Areas.* Aldershot, England: Ashgate Publishing, Ltd; 2005. p. 33–45.
- 827 65. Nunnally JC, Bernstein IH. *Psychometric Theory.* 3rd ed. New York, NY: McGraw-Hill,
828 Inc.; 1994.
- 829 66. Sainani KL. Introduction to Principal Components Analysis. *PM&R.* 2014 Mar;6(3): 275-
830 278.
- 831 67. Everitt B, Hothorn T. *An Introduction to Applied Multivariate Analysis with R.* 1st ed.
832 Gentleman R, Hornik K, Parmigiani G, editors. Vol. 1. New York: Springer; 2011.
- 833 68. Kassambara A. *Practical Guide to Principal Component Methods in R (Multivariate
834 Analysis).* 1st ed. Vol. 1. *Statistical Tools for High-throughput Data Analysis (STHDA)*
835 (<http://www.sthda.com>); 2017.

- 836 69. Kass RE, Steffey D. Approximate Bayesian Inference in Conditionally Independent
837 Hierarchical Models (Parametric Empirical Bayes Models). *J Am Stat Assoc* [Internet].
838 1989 Sep 1;84(407):717–26. doi: 10.1080/01621459.1989.10478825
- 839 70. Casella G. An Introduction to Empirical Bayes Data Analysis. *Am Stat* [Internet]. 1985
840 May 1;39(2):83–7. doi: 10.1080/00031305.1985.10479400
- 841 71. Savic RM, Karlsson MO. Importance of Shrinkage in Empirical Bayes Estimates for
842 Diagnostics: Problems and Solutions. *AAPS J* [Internet]. 2009;11(3):558–69. doi:
843 10.1208/s12248-009-9133-0
- 844 72. Abseysekera W SM. Use of Schoenfeld’s global test to test the proportional hazards
845 assumption in the Cox proportional hazards model: an application to a clinical study. *J*
846 *Natil Sci Found Sri Lanka*. 2009;37(1):41–51.
- 847 73. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and
848 guidance for practice. *Stat Med*. 2011 Feb 20;30(4):377–99.
- 849 74. Masanja H, Savigny D de, Smithson P, Schellenberg AJ, John T, Mshinda H, et al. Child
850 survival gains in Tanzania: analysis of data from demographic and health surveys. *The*
851 *Lancet*. 2008;371(9620):1276–1283.
- 852 75. Singh NS, Huicho L, Afnan-Holmes H, John T, Moran AC, Colbourn T, et al. Countdown
853 to 2015 country case studies: systematic tools to address the “black box” of health
854 systems and policy assessment. *BMC Public Health*. 2016 Sep;16 Suppl 2(Suppl 2):790.
- 855 76. Leonard KL, Mliga GR, Haile Mariam D. Bypassing Health Centres in Tanzania:
856 Revealed Preferences for Quality. *J Afr Econ* [Internet]. 2002 Dec 1;11(4):441–71. doi:
857 10.1093/jae/11.4.441
- 858 77. Leonard KL, Masatu MC. Variations In The Quality Of Care Accessible To Rural
859 Communities In Tanzania. *Health Aff*. 2007 Jan;26(Suppl2):w380–92.
- 860 78. Kahabuka C, Kvåle G, Moland KM, Hinderaker SG. Why caretakers bypass Primary
861 Health Care facilities for child care-a case from rural Tanzania. *BMC Health Serv Res*.
862 2011;11(1):1–10.
- 863 79. Kanté AM, Exavery A, Phillips JF, Jackson EF. Why women bypass front-line health
864 facility services in pursuit of obstetric care provided elsewhere: a case study in three rural
865 districts of Tanzania. *Tropical Medicine & International Health*. 2016 Apr;21(4):504–14.
- 866 80. Serván-Mori E, Contreras-Loya D, Gomez-Dantés O, Nigenda G, Sosa-Rubí SG, Lozano
867 R. Use of performance metrics for the measurement of universal coverage for maternal
868 care in Mexico. *Health Policy Plan*. 2017;32(5):625–33.
- 869 81. Kruk ME, Leslie HH, Verguet S, Mbaruku GM, Adanu RMK, Langer A. Quality of basic
870 maternal care functions in health facilities of five African countries: an analysis of
871 national health system surveys. *Lancet Glob Health*. 2016 Nov;4(11):e845–55.

- 872 82. Nguhiu PK, Barasa EW, Chuma J. Determining the effective coverage of maternal and
873 child health services in Kenya, using demographic and health survey data sets: tracking
874 progress towards universal health coverage. *Tropical Medicine & International Health*.
875 2017 Apr;22(4):442–53.
- 876 83. Sharker MA, Nasser M, Abedin J, Arnold BF, Luby SP. The risk of misclassifying
877 subjects within principal component based asset index. *Emerg Themes Epidemiol*.
878 2014;11(1):1–8.
- 879

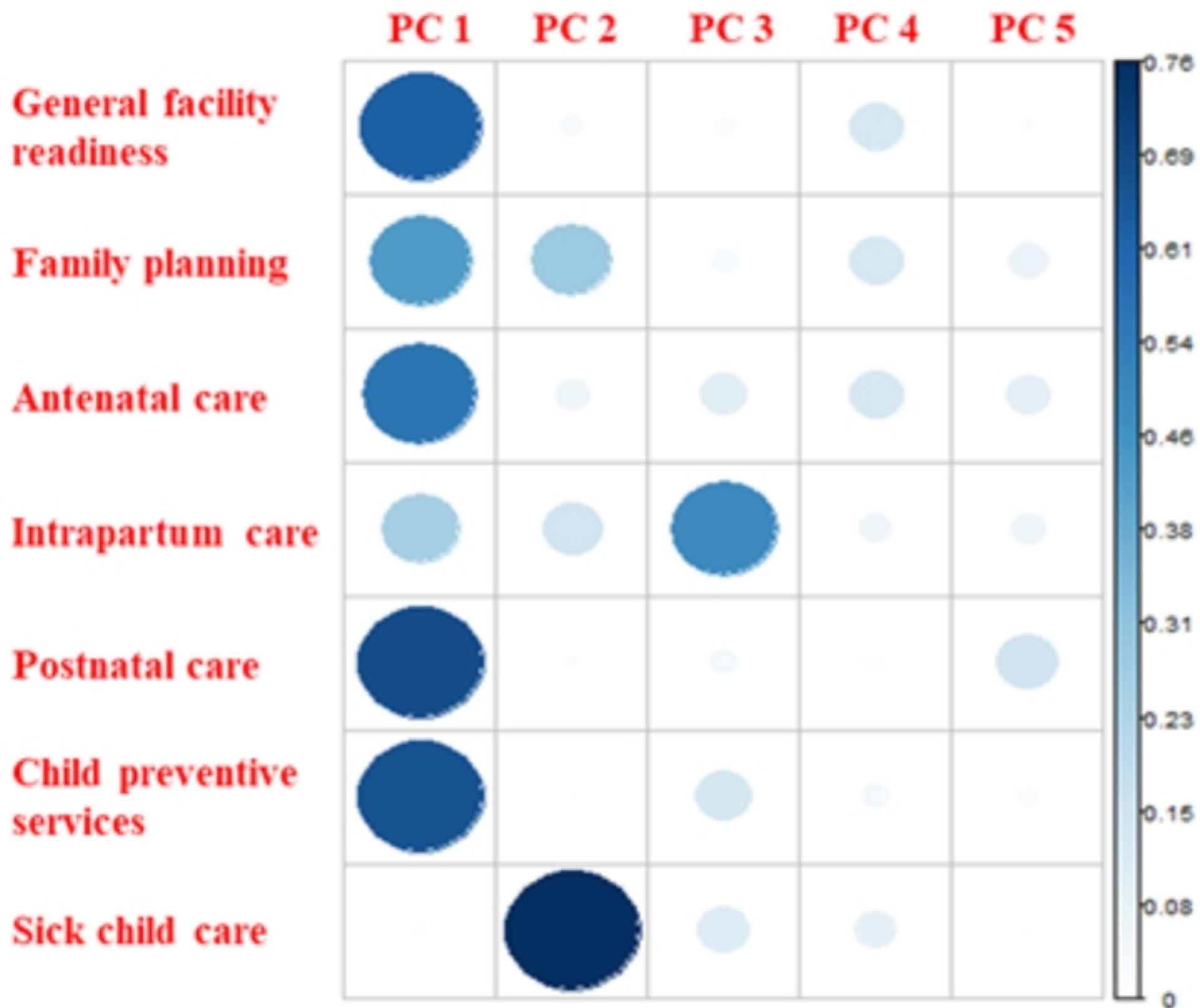
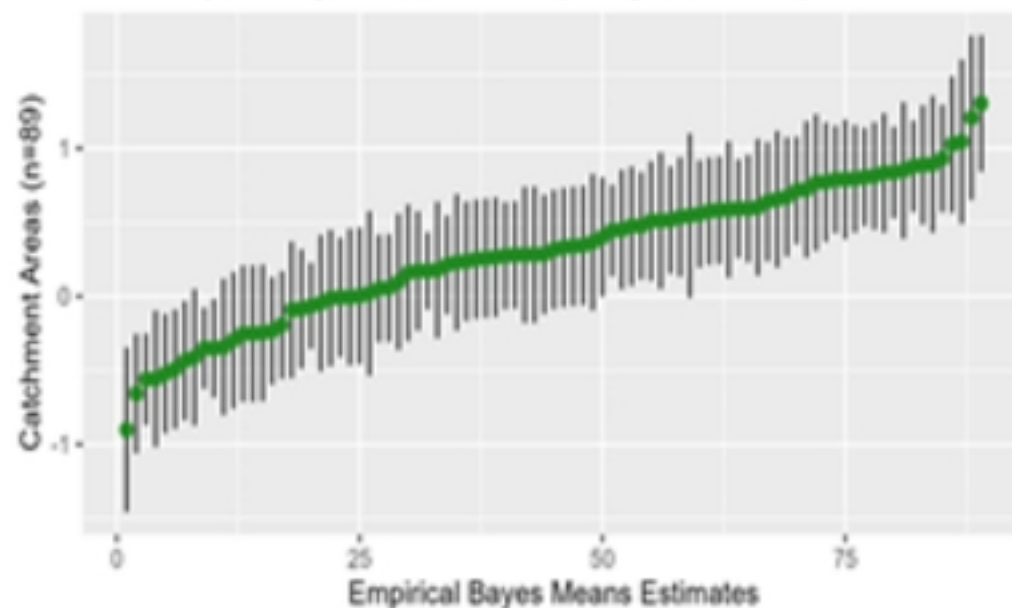


Figure 3b

Distribution of Implementation Strength Scores - Scale 1

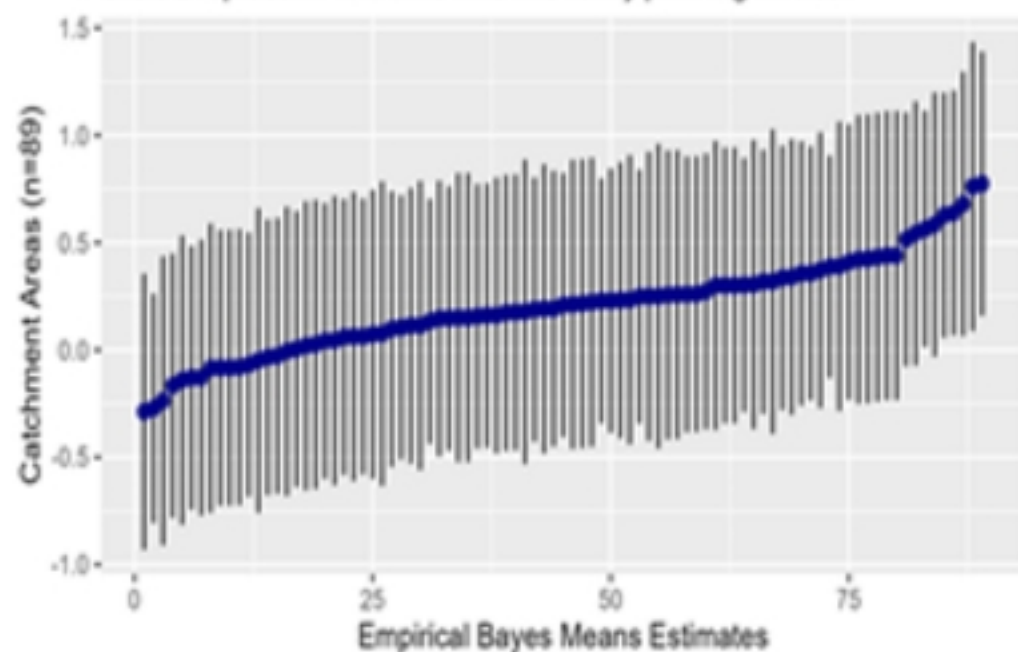
Scale 1 represents preventive services (mostly child related)



Green points are Empirical Bayes Means Estimates, black lines are 95% confidence intervals

Distribution of Implementation Strength Scores - Scale 2

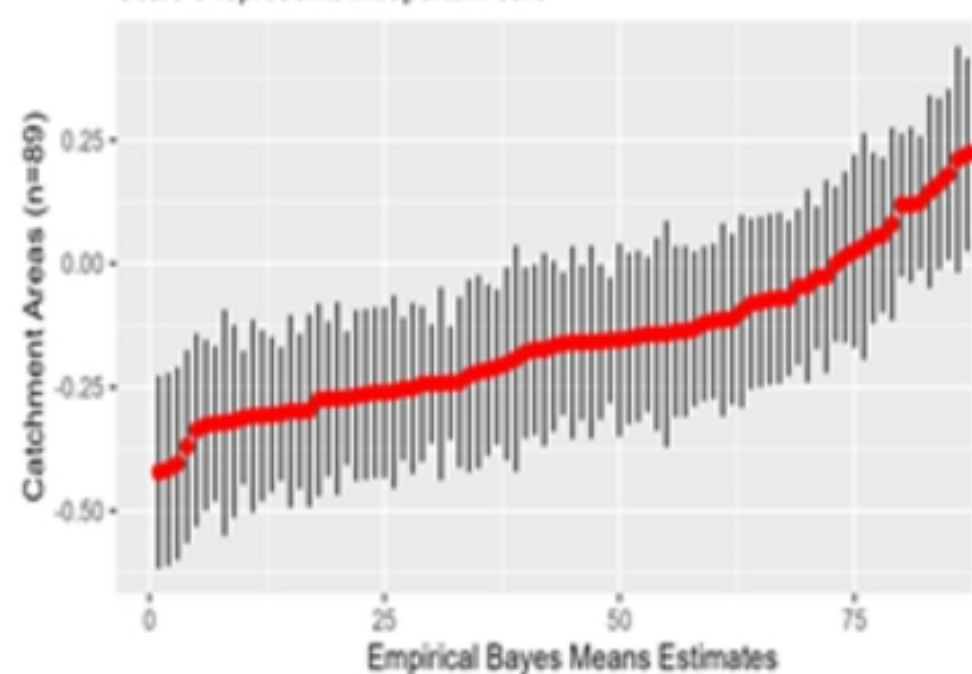
Scale 2 represents sick child care and family planning services



Blue points are Empirical Bayes Means Estimates, black lines are 95% confidence intervals

Distribution of Implementation Strength Scores - Scale 3

Scale 3 represents intrapartum care



Red points are Empirical Bayes Means Estimates, black lines are 95% confidence intervals

Figure 4

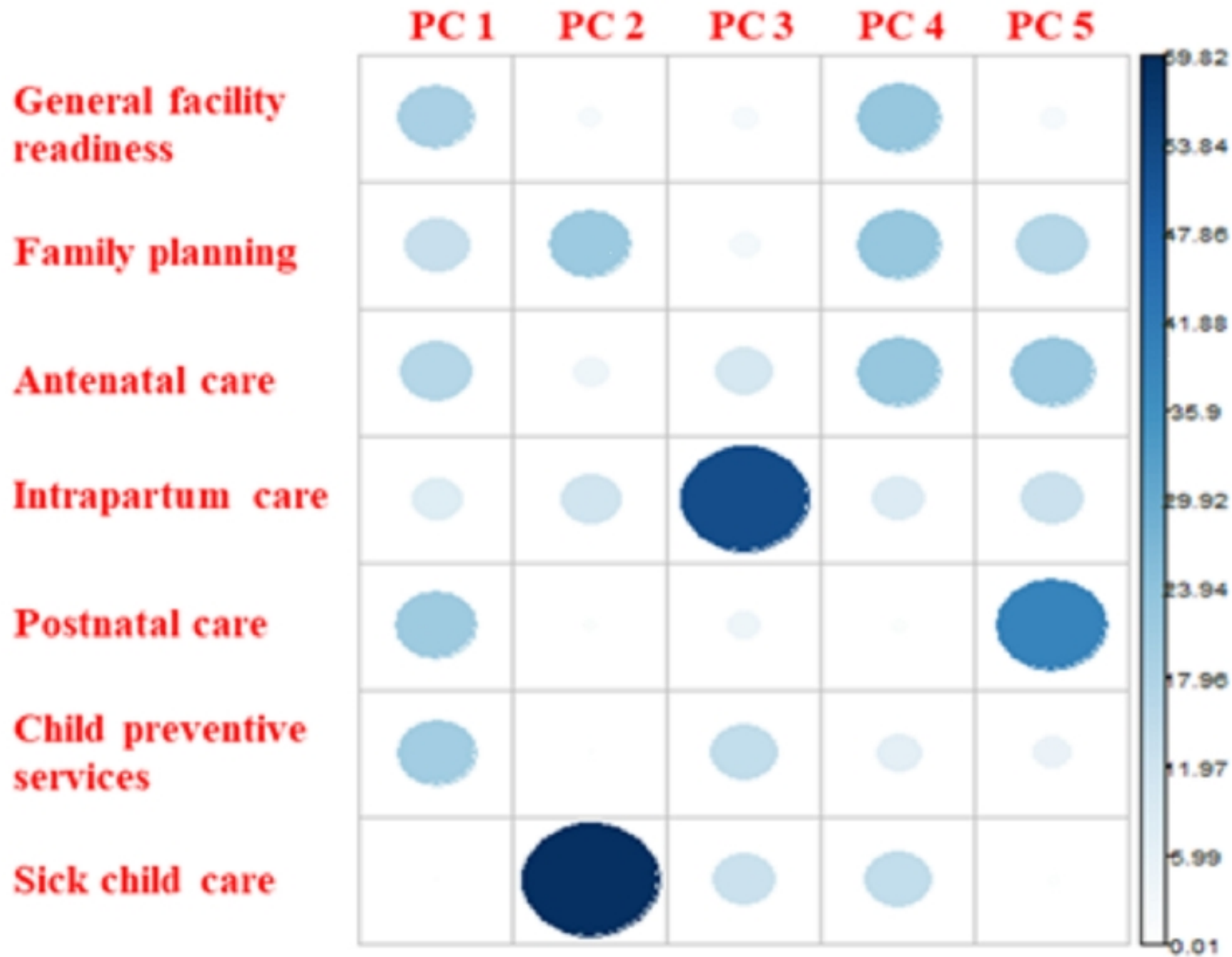


Figure 3a

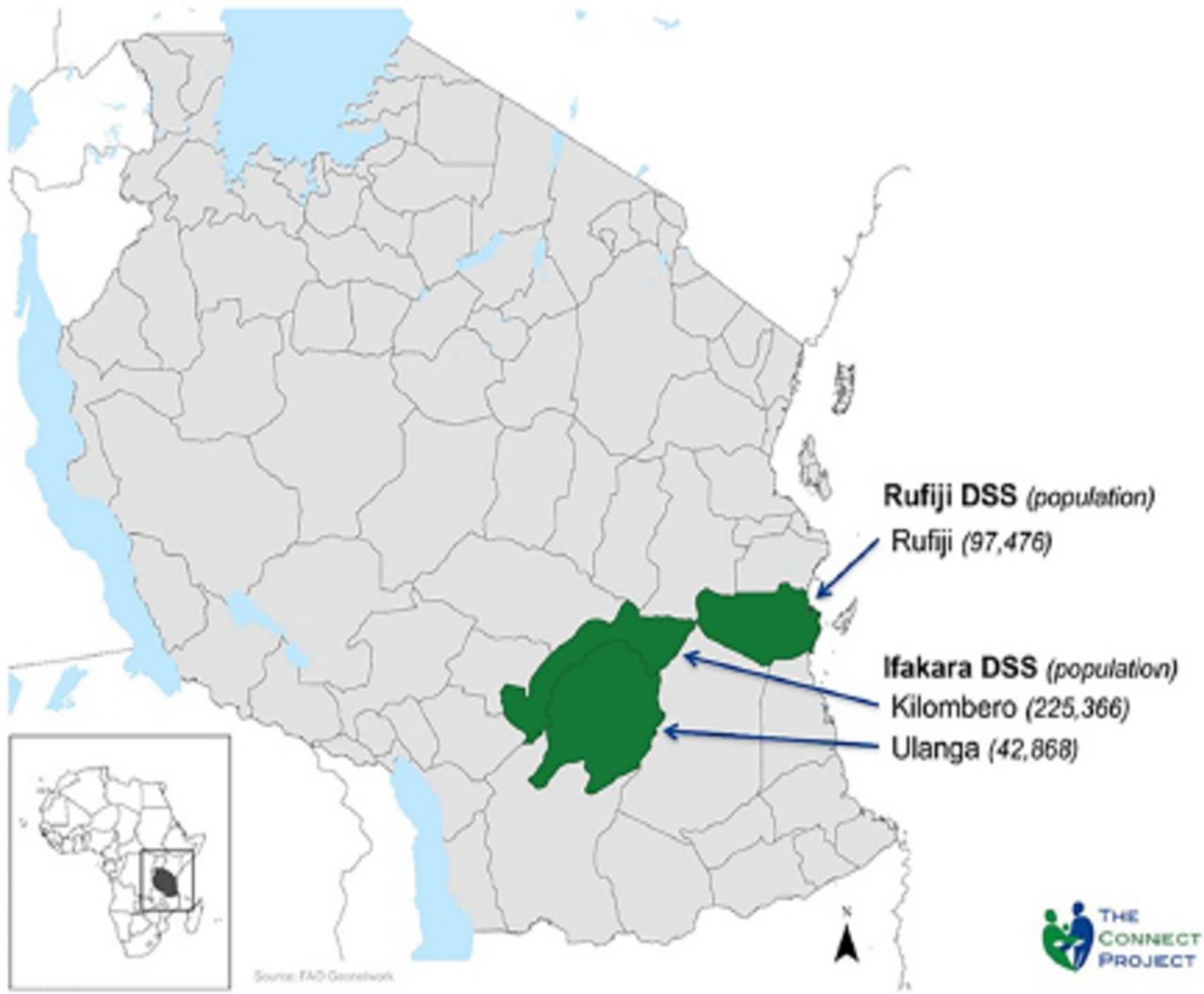
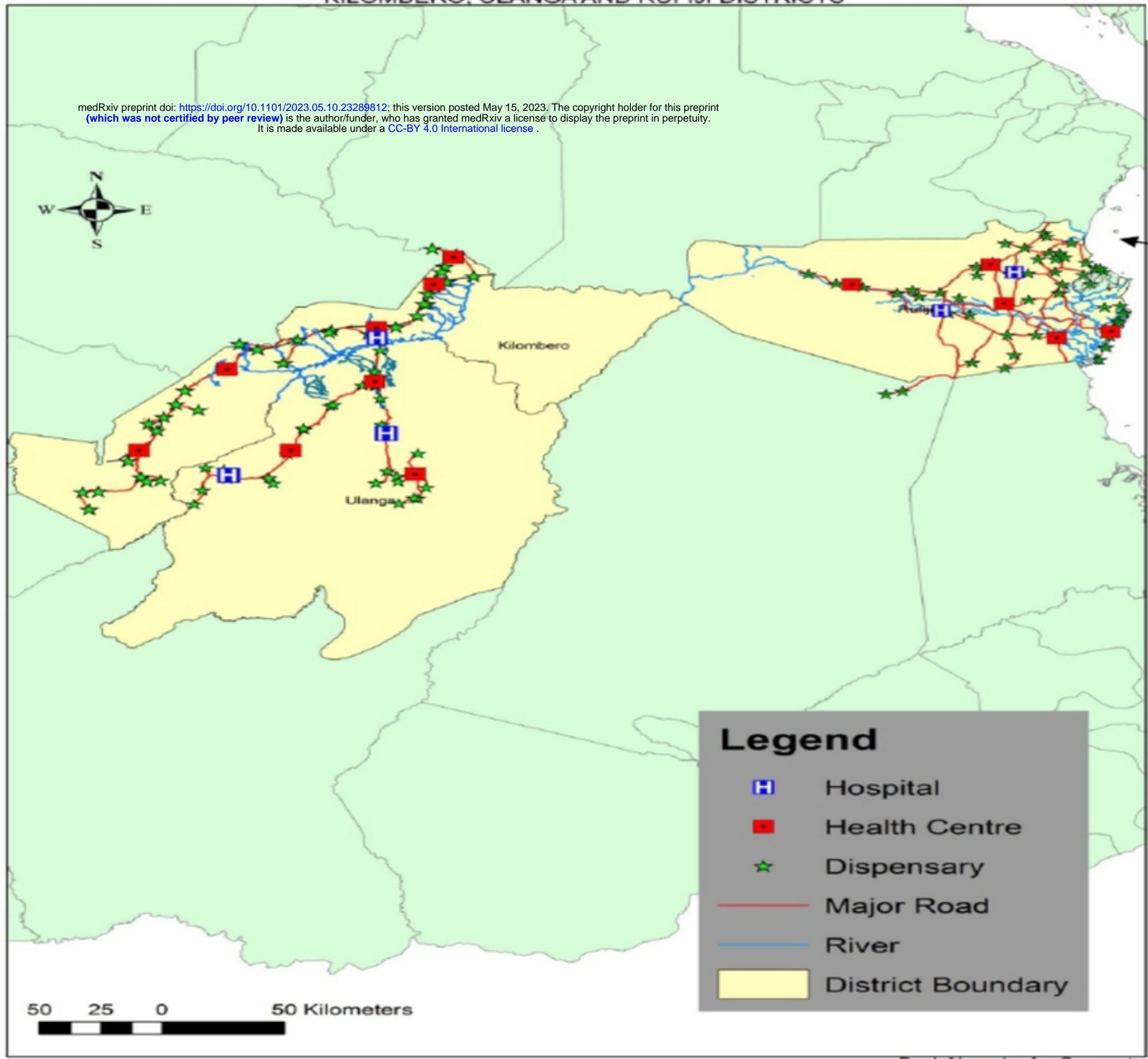


Figure 1



medRxiv preprint doi: <https://doi.org/10.1101/2023.05.10.23289812>; this version posted May 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).



By: J. Noronha for Connect

Figure 2