

1 **Comparison of causes of stillbirth and child deaths as determined by verbal**
2 **autopsy and minimally invasive tissue sampling**

3
4 Nega Assefa^{1*}, Anthony Scott², Lola Madrid², Merga Dheresa¹, Gezahegn Mengesha¹,
5 Shabir Mahdi³, Sana Mahtab³, Ziyaad Dangor³, Nellie Myburgh⁴, Lesego Kamogelo
6 Mothibi⁴, Samba O. Sow⁴, Karen L. Kotloff⁵, Milagritos D. Tapia⁵, Uma U. Onwuchekwa⁴,
7 Mahamane Djiteye⁴, Rosauero Varo^{6,7}, Inacio Mandomando^{7,8}, Ariel Nhacolo⁷, Charfudin
8 Sacoor⁷, Elisio Xerinda⁷, Ikechukwu Ogbuanu⁹, Solomon Samura¹⁰, Babatunde Duduyemi¹¹,
9 Alim Swaray-Deen¹², Abdulai Bah¹³, Shams El Arifeen¹⁴, Emily S Gurley¹⁵, Mohammed
10 Zahid Hossain¹⁶, Afruna Rahman¹⁷, Atique Iqbal Chowdhury¹⁶, Quique, Bassat^{6,7,8}, Portia
11 Mutevedzi¹⁸, Argeseanu Solveig¹⁹, Dianna Blau¹⁹□, Cyndy Whitney¹⁸□

12
13 ¹ *College of Health and Medical Sciences, Haramaya University, Harar, Ethiopia*

14 ² *London School of Hygiene and Tropical Medicine, London, United Kingdom*

15 ³ *South African Medical Research Council Vaccines and Infectious Diseases Analytics*
16 *Research Unit, University of the Witwatersrand, Johannesburg, South Africa*

17 ⁴ *Centre pour le Développement des Vaccins (CVD-Mali), Ministère de la Santé, Bamako,*
18 *Mali;*

19 ⁵ *Department of Pediatrics, Center for Vaccine Development and Global Health, University*
20 *of Maryland School of Medicine, Baltimore, Maryland, USA*

21 ⁶ *ISGlobal, Hospital Clínic, Universitat de Barcelona, Barcelona, Spain*

22 ⁷ *Centro de Investigação em Saúde de Manhiça (CISM), Maputo, Mozambique*

23 ⁸ *Instituto Nacional de Saude, Ministerio de Saude, Maputo, Mozambique*

24 ^{NOTE:} This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Crown Agents, Freetown, Sierra Leone

25 ¹⁰ *World hope international, Makeni, Sierra Leone*

26 ¹¹ *University of Sierra Leone Teaching Hospitals Complex, Sierra Leone*

27 ¹² *University of Ghana Medical School*

28 ¹³ *FOCUS 10000, Freetown, Sierra Leone*

29 ¹⁴ *Maternal and Child Health Division, icddr,b, Dhaka, Bangladesh*

30 ¹⁵ *Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland,*
31 *United States of America*

32 ¹⁶ *International Centre for Diarrhoeal Disease Research Bangladesh (icddr,b), Dhaka,*
33 *Bangladesh*

34 ¹⁷ *PEI, Infectious Disease Division, icddr,b, Dhaka, Bangladesh*

35 ¹⁸ *Global Health Institute, Emory University, Atlanta, Georgia, United States of America*

36 ¹⁹ *Global Health Center, Centers for Disease Control and Prevention, Atlanta, Georgia,*
37 *United States of America*

38 #a **Corresponding author:** Nega Assefa, College of Health and Medical
39 Sciences, Haramaya University, Harar Campus, SGS building, Ethiopia
40 nassefa@hararghe.org

41 These authors contributed equally to this work

42

43 **Keywords:** Cause of death, Determination of Cause of Death, CHAMPS, sub-Saharan Africa,
44 Open VA, InterVA-4 package

1 **Abstract**

2 **Background:** In resource-limited settings where vital registration and medical death
3 certificates are unavailable or incomplete, verbal autopsy (VA) is often used to attribute causes
4 of death (CoD), identify the distribution and trends of diseases, and prioritize resource
5 allocation and interventions. However, VA findings can be non-specific, as this tool is based
6 on family members' recall of symptoms rather than objective diagnostic testing. We aimed to
7 compare the CoD diagnoses obtained in stillbirths and children below five years of age (<5s)
8 through two very different approaches; namely: 1) VA; and 2) the results obtained through the
9 use of Minimally Invasive Tissue Sampling (MITS) and rigorous diagnostic testing, as part of
10 the approach proposed by the Child Health and Mortality Prevention Surveillance (CHAMPS).

11 **Methods:** CHAMPS identified stillbirths and deceased children <5s in real time between 2017
12 and 2021 in catchment areas in seven low- and middle-income countries (LMICs): Bangladesh,
13 Ethiopia, Kenya, Mali, Mozambique, Sierra Leone, and South Africa. Deaths were eligible for
14 MITS if identified <24 hours after death, legal concerns were not present, burial had not
15 occurred, and parents consented. CHAMPS teams utilized information from MITS and VA to
16 determine the causes of death (CoDs); if not eligible for MITS, the InterVA software utilized
17 only VA information to determine the CoDs. CHAMPS attributed CoD using expert panels
18 that reviewed clinical evidence microbiological, and histopathological results from MITS to
19 derive the CoDs (Determination of Cause of Death [DeCoDe]). The InterVA4 package of
20 OpenVA software automatically assigned the underlying CoDs using the Bayesian
21 probabilistic modeling technique. These automatically assigned CoDs from OpenVA were
22 compared to the gold-standard of the CHAMPS-attributed CoDs to evaluate both systems'
23 agreement, weaknesses, and strengths using Lin's concordance correlation coefficient.

24 **Results:** Data from 2852 deaths that underwent MITS were analysed. The most common age
25 categories were stillbirths (n=1075, 37.7%) and neonatal deaths (n=1077, 37.8%). Overall
26 concordance of InterVA4 and DeCoDe in assigning causes of death across surveillance sites,
27 age groups, and causes of death was poor (0.75 with 95% CI: 0.73 – 0.76) and lacked precision.
28 We found substantial differences in agreement among surveillance sites, with Mali showing
29 the lowest and Mozambique and Ethiopia the highest concordance. Lin's concordance
30 correlation coefficient for children aged < 1 year was 0.69 (95%CI: 0.65 – 0.71), and for
31 children aged 1-4 years was 0.28 (95%CI: 0.19 – 0.37)

32 **Conclusion:** The InterVA4 assigned CoD agrees poorly in assigning causes of death for under-
33 fives and stillbirths. Because VA methods are relatively easy to implement, such systems could
34 be more useful if algorithms were improved to more accurately reflect causes of death, for
35 example, by calibrating algorithms to information from programs that used detailed diagnostic
36 testing to improve the accuracy of COD determination.

37 **Introduction**

38 In contrast to countries in high-income countries, most low and middle-income
39 countries (LMIC) that have higher rates of mortality among children lack adequate systematic
40 mortality surveillance [1]. For example, death registration coverage varies from nearly 100%
41 in the WHO European region to less than 10% in the WHO African region [2]. In LMICs,
42 deaths are often not attended by health professionals, not medically certified, not recorded in
43 a timely way, and, even when recorded, the information is stored inappropriately [3]. LMICs
44 also do not have the infrastructure or resources to establish and maintain data systems that
45 conclusively identify causes of death in their populations [4]. Not having appropriate
46 legislation or health policies on data systems compounds these challenges, leading to
47 ineffective formulation and implementation of interventions to reduce mortality at a
48 population level [5, 6].

49 One relatively simple method to try to identify the CoDs is through a verbal autopsy
50 (VA) [7-10]. To conduct a VA, workers who are trained on the method but who typically lack
51 specific clinical training interview family members or caregivers of the deceased using a
52 structured questionnaire; the workers also solicit a complete narrative of the circumstances
53 surrounding death [11]. Causes of death (CoDs) can then be easily generated from the
54 structured questionnaire responses by probabilistic analytic algorithms that are freely
55 available and accessible online. Because other modalities, such as physician coding, may not
56 be efficient, affordable, and sustainable in resource-limited settings, publicly available CoDs
57 generating software from the VA has become an essential public health tool for mortality
58 estimation and identifying population-level CoDs in resource-limited settings [12, 13].

59

60 However, VA has several weaknesses [14, 15]. The method can produce conflicting
61 and unreliable CoD results because it relies on the quality and accuracy of information

62 provided by family members who typically lack clinical training. The community's
63 sociocultural and recall biases can affect families' responses. The VA forms do not collect
64 information on known or pre-existing medical conditions determined based on diagnostic
65 testing, as the families are the proxy respondents who might not have access to past clinical
66 information [16-18]. The presence of multiple VA algorithms and the tool's inherent
67 limitations with accurately determining pre-existing or new medical problems make it
68 challenging to understand and determine the CoD for conditions with complex cause-of-death
69 pathways or highly non-specific signs and symptoms [19, 20]. In addition, the VA does not
70 generate the complete mortality pathways, such as the immediate or morbid pathways, but
71 only determines the underlying CoDs with various probabilistic scores, which may not
72 adequately capture complicated medical histories. For stillbirths, the VA only describes the
73 body's condition, which has been shown to not accurately reflect or even impossible to
74 determine the causes of death among stillbirths [21, 22]

75 The Child Health and Mortality Prevention Surveillance (CHAMPS) is a collaborative
76 network in sub-Saharan Africa and South Asia that uses other additional approaches,
77 including Minimally Invasive Tissue Sampling (MITS), clinical data, histopathological and
78 microbiological findings, and the VA narrative itself, to provide reliable, detailed, and
79 specific causes of stillbirths and child deaths [23-25]. With the support of the Bill & Melinda
80 Gates Foundation, CHAMPS was launched in 2017 in several high child-mortality countries
81 to provide reliable data on cause-specific mortality, which is fundamental to evidence-based
82 health policy and public health interventions [26]. CHAMPS uses thorough postmortem
83 diagnostic testing along with some parts of the VA (open narrative and raw answers to VA
84 questionnaire, but not the VA-derived diagnoses) and review of clinical records. A local
85 panel of experts reviews all the information package and assigns underlying, intermediate,
86 and immediate causes of death, a process called Determination of Cause of Death (DeCoDe).

87 While CHAMPS' methods produce high-quality cause-of-death information for those
88 children evaluated, the postmortem diagnostic testing protocols require rapid death
89 identification and collection of specimens before a child is buried. Therefore, many deaths in
90 CHAMPS catchment areas do not undergo such testing; in contrast, evaluation of deaths
91 using VA can be done after burial, at a family's convenience, which is usually 2 to 12 weeks
92 [27].

93 This study compares the type, quality, and amount of CoD information generated
94 from "VA only" and CHAMPS's methods; we also assess the concordance of the results
95 generated from the two approaches. While CHAMPS generates specific microbiology and
96 pathology diagnoses, this study focuses on the accuracy of the CoDs assigned by both
97 systems. Methodologically speaking, although VA data was taken using the WHO 2016
98 questionnaire, the ICD-10 diagnosis determined by the CHAMPS methods was mapped to
99 syndromic categories from the 2016 WHO Verbal Autopsy guideline for comparative
100 purposes. By analysing data from CHAMPS sites in seven countries in Sub-Saharan Africa
101 and Southeast Asia (Bangladesh, Ethiopia, Kenya, Mali, Mozambique, Sierra Leone, and
102 South Africa), we aimed to inform public health leaders and policymakers of the strengths
103 and weaknesses, relevance, and consistency of these approaches.

105 **Methods**

106 *Study settings and design*

107 The CHAMPS Network longitudinally collects robust and standardized data from its
108 sites to understand and track preventable causes of childhood death in high-mortality areas.
109 The CHAMPS network details have been published elsewhere. [28-30]. All CHAMPS
110 network sites are in research centers with pre-existing Demographic Health Surveillance
111 Systems (HDSS) or have built capacity to closely follow up their catchment's population,
112 enabling them to capture periodic mortality surveillance data. An HDSS is an open, dynamic
113 cohort that follows residents of a geographically defined area over time, tracking the
114 occurrence of births, deaths, marriages, pregnancies, and migrations. HDSS teams in
115 CHAMPS sites enumerate these events during routine household visits. To identify stillbirths
116 and deaths in children under five as soon as they occur, mortality surveillance in CHAMPS
117 sites also involves community informants, healthcare workers, and other methods.

118 *Data collection*

119 CHAMPS study procedures have been published elsewhere [31]. Briefly, data from deaths
120 identified through HDSS and mortality surveillance are collected longitudinally from notified
121 deaths in the communities and health facilities within the catchment. An <5s or stillbirth
122 identified within 24 hours of death or 72 hours if refrigerated whose family had been living in
123 the respective catchment area for at least four to six months is eligible to be enrolled for the
124 CHAMPS and requested to provide consent for Minimally Invasive Tissue Sampling (MITS).
125 MITS includes postmortem collection of swabs, postmortem biopsies of vital organs, and
126 body fluids for cytopathologic and microbiologic examination. Clinical information found at
127 the health facilities and the community where the stillbirths and <5s deaths occurred is also
128 collected, and families of the deceased are interviewed using the VA questionnaire.

129 ***Verbal autopsy questionnaire***

130 All sites use the WHO-2016 VA questionnaire, customized to include content
131 enhancements, skip logic, and unit of measurement corrections for the CHAMPS study [32].
132 VA questionnaires were translated into local languages and collected information on age, sex,
133 place of death, and symptoms observed during the late-life period of the deceased. The
134 questionnaire also contains the symptom duration checklist, which is arranged loosely around
135 anatomical systems and is intended to be informative in leading to a diagnosis of probable
136 CoDs and narrowing the number of possible differential diagnoses.

137 ***Causes of Death assignment from VA***

138 We used the InterVA-4 package from Open-VA to auto-generate the cause for each
139 enrolled death at the surveillance site [33, 34]. Open-VA uses Bayesian probabilistic
140 modeling to assign likelihoods to causes of death based on coded responses to verbal autopsy
141 questionnaires and ascribes corresponding ICD-10 codes [35, 36]; InterVA-4 algorithms do
142 not consider information in the narrative section of the VA. This system mainly generates one
143 likely CoDs and, if a single cause is not clear, three causes with probability values. The
144 generated CoD with the highest probability was considered the underlying cause for
145 comparison with the CoD assigned by CHAMPS DeCoDe.

146 ***Determination of Cause-of-Death (DeCoDe) using Minimally Invasive Tissue***
147 ***Sampling (MITS)***

148 Following the World Health Organization (WHO) application of the International
149 Classification of Diseases – Version 10 (ICD-10), the DeCoDe expert panel determined the
150 underlying cause and, for some deaths, one or more intermediate causes and an immediate
151 CoD [35, 37]. Because we also compared the immediate CoDs assigned by the DeCoDe with
152 the InterVA4’s underlying CoDs, as the InterVA4 does not designate immediate CoDs as the

153 DeCoDe and found no significant difference, we decided Only to use the underlying CoDs
154 assigned by DeCoDe for comparison purposes.
155 To ensure the DeCoDe panel is aligned with the correct diagnoses and decrease bias, all panel
156 members of each surveillance site follow a standard operating procedure and CHAMPS
157 Diagnosis Standards [38].
158 The assigned causes of death by the DeCoDe panel were converted and categorized to the
159 corresponding VA diagnosis using the 2016 WHO VA category definitions of the verbal
160 autopsy standard [39]. The standard has a conversion table that shows and defines the VA
161 diagnosis category and title with its corresponding ICD-10 codes. This conversion and
162 categorization enable comparison of the generated CoD InterVA4 with the DeCoDe, which is
163 considered a gold standard for concordance and accuracy.

164 ***Ethical clearance***

165 Ethical clearances from the respective institutions and national ethical clearance
166 bodies have been secured for HDSS and CHAMPS activities. HDSS activities have standing
167 approvals for continuing routine activities, including VA. All participants provided informed,
168 voluntary, written consent. Consent was obtained from the responsible person in the family
169 (the head of the household, the mother of the deceased child, or any eligible family member).
170 Written informed consent was obtained from the parent/guardian of each participant under 18
171 years of age. To keep anonymity and confidentiality, we did not share data that contained
172 participants' personal identifiers with any third party.

173 ***Quality control***

174 Individuals who completed at least a high school education and had experience
175 working in the existing HDSS collected VA data. They received a two-week training on the
176 questionnaires, recording, contacting close relatives, and data collection procedures. The

177 training included sessions on discussing individual symptoms and their description in the
178 local language for easy recognition by the respondents and demonstration of interviewing
179 techniques by research team members. The field coordinators and supervisors continuously
180 monitor data collection in the field to check progress and resolve problems that enumerators
181 may have encountered during fieldwork.

182 *Data management and analysis*

183 Data were analysed using STATA version 16. Means and standard deviations (SDs)
184 were presented for continuous variables, medians and interquartile ranges for skewed
185 variables, and counts and percentages for categorical variables. Demographic characteristics
186 included age, gender, occupation, religion, and household size. Variables with more than
187 45% missing data were excluded.

188 We considered stillbirth as the absence of life after the
189 viability of pregnancy (≥ 28 weeks of gestation) and before and
190 during the baby's delivery. Neonatal death was defined as a death
191 in a live-born baby in the first 28 days of life. We further
192 classified neonatal death into very early, early, and late neonatal
193 death if the death occurred in the first 24 hours (day 0), 1–6 days,
194 and 7 to 28 days, respectively [40]. Infant death was defined as a
195 baby's death after 28 days of life and before the first birthday, and
196 child death as death before celebrating his/her 5th year birthday [41].

197 Cause-specific mortality fractions (CSMF) for each surveillance site and CoD were
198 computed by dividing the number of deaths due to specific causes assigned by either
199 InterVA-4 and CHAMPS's DeCoDe over the total number of deaths evaluated. The

200 underlying causes of death from InterVA and DeCoDe were compared for agreement and
201 pattern in assigning the diagnosis.

202 After the respective underlying causes of death that DeCoDe assigned were mapped
203 and matched to its corresponding verbal autopsy standard, the agreement of both methods
204 was evaluated using their concordance and accuracy of CSMF. We compared the CSMF of
205 InterVA4 against DeCoDe using Lin's Concordance Correlation Coefficient (LCC), [42]
206 which was calculated using a user-defined command made for Stata – “Concord”[43].

207 The LCC determines how far the observed data deviate from the line of perfect
208 concordance, a line at 45 degrees in a scatterplot. Lin's coefficient increases in value as a
209 function of the nearness of the data's reduced major axis to the line of perfect concordance
210 (the accuracy of the data) and of the tightness of the data about its reduced major axis (the
211 precision of the data). The bias bias-correction factor shows how far the best-line of shift is
212 from the perfect concordance. The program (“Concord”) produces the LCC by multiplying
213 the “Pearson correlation coefficient, r ” with the bias-correction factor. Whereas the “Pearson
214 correlation coefficient, r ” is the measure of precision, the bias-correction factor is for
215 accuracy [43].

216 The LCC was stratified across surveillance sites, age classification, and enrolment
217 location to evaluate the performance of InterVA4. The stratification of the group was
218 according to the WHO 2016 VA instrument guideline [36]: children aged < 1 year and aged
219 1-4 years. Accuracy is the measurement of the validity of a measurement's exact value or
220 how close the predicted value obtained in data is to the true value. Precision is defined as the
221 degree of reproducibility of using the same measurement or procedure to measure the degree
222 of consistency of independent measurements of the same variable [44]. The interpretation of
223 the LCC we used is < 0.8 is poor, 0.81-90 –as good, and > 0.9 is excellent [45]. We also used
224 the same interpretation for accuracy and precision.

225 Furthermore, to complement LCC in measuring the agreement between InterVA4 and
226 DeCoDe, a mortality fraction ratio was calculated by dividing the CSMF generated by the
227 InterVA4 with the DeCoDe's (InterVA4 CSMF/DeCoDe CSMF) by surveillance site and for
228 specific CoDs at a 95% confidence interval generated using the Koopman method to identify
229 whether the interpretation between the two methods was lower or higher than expected [46].
230 This statistical method produces "the Koopman asymptotic score interval" for the ratio of
231 probabilities in two-by-two contingency tables and works well for small sample sizes. The
232 purpose of calculating these CIs was not to demonstrate statistical significance but to identify
233 whether the CSMF ratio between InterVA-4 and DeCoDe interpretations was significantly
234 lower or higher than that expected from chance, considering the number of cases involved.

235 **Results**

236 CHAMPS sites identified 7221 unique deaths (including stillbirths), of which 6,909 (95.7%)
237 were enrolled from February 1, 2017, through December 30, 2021 (Figure 1). Of 6,909
238 enrolled deaths, 338 (4.9%) observations were removed from the analysis because they were
239 missing CoDs generated from the InterVA-4 package of the Open-VA because of
240 transcription errors, and 77 were removed because of a conflicting date of birth or death and
241 CoDs. These deaths were also removed from the analysis. Of the remaining 6494 deaths,
242 2340 (36.0%) were stillbirths, 2321 (35.7%) were neonates, 967 (14.9%) were infants, and
243 866 (13.3%) were children aged 1-<5 years. Of these, 3641 deaths were excluded as they
244 were not enrolled for MITS and only had InterVA-generated CoD. Therefore, we analyzed
245 2853 (43.9%) of 6494 deaths enrolled for MITS and subsequently had CoD information
246 generated from both DeCoDe and VA.

247 Of 2853 eligible deaths where both MITS and VA were undertaken, 1075 (37.7%)
248 were stillbirths, 1077 (37.8%) neonatal deaths, 365 (12.8%) infant deaths, and 336 (11.9%)
249 deaths of children aged 1-4 years. Around 30% (654) were enrolled from South Africa, 19%
250 (545) from Kenya, 16.7% (476) from Mozambique, 12.2% (348) from Bangladesh, 11% from
251 Sierra Leone (316) and Ethiopia (311), and 7% (203) from Mali. Across sites, the DeCoDe
252 panel could not determine the CoDs for 78 (2.7%) of all deaths enrolled; of these, about half
253 (40, 51.3%) were stillbirths, 15 (19.2%) were neonatal deaths, 15 (19.2%) infants, and 8
254 (10.3%) children aged 1-4 deaths (Table 1).

255

256

257 ***Characteristics of study population and mortality groups across sites***

258 Across sites, the mean age at death for U5 and newborns was 4 ± 10 months; females
259 accounted for 44.1% of deaths (1251/2853), and more than half (55.1%) occurred in the dry
260 season. The mean age of death was $4 (\pm 5.4)$ days for neonates, $5.4 (\pm 4)$ months for infants,
261 and $2.2 (\pm 1.1)$ years. The mean gestational age for stillbirths was 34.3 weeks (95% CI 33.3,
262 35.6 weeks). Of 1075 recorded stillbirths, (207 (19.3%) were in Mozambique, 199 (18.5%) in
263 Ethiopia, 177 (16.5%) in Bangladesh, 160 (14.9%) in Kenya, 153 (14.2%) in South Africa,
264 99 (9.2%) in Sierra Leone, and 80 (7.4%) in Mali. Overall, a large majority (89%) of deaths
265 occurred in health facilities. About one-third of infant (111/365, 30.4%) and child deaths
266 (102/336, 30.4%) occurred in the community. However, nearly all enrolled stillbirths
267 (1037/1075, 96.5%) and neonatal deaths (1015/1077, 92.4%) occurred in health facilities
268 (**Tables 1 and 2**).

269 The overall concordance of diagnoses across the surveillance sites and age groups was
270 0.75 (Table 2). The interVA4 method of assigning CoDs had better accuracy, but its precision
271 compared to the DeCoDe was poor (<0.8). Stratified by surveillance sites, the overall
272 concordance of all <5 s deaths was lowest in Mali (0.64), and Ethiopia (0.83) and
273 Mozambique sites (0.84) had good overall concordance.

274 **Figure 2** shows the overall LCC of the CSMF generated by the InterVA4 against
275 DeCoDe's underlying causes of the death. We found the overall concordance level between
276 the two systems to be poor, 0.75 (95%CI 0.73 – 0.76). The precision of the concordance was
277 0.98, while the accuracy was 0.76. The concordance coefficients were nearly the same across
278 sexes for all CoDs and were higher for <5 s enrolled at health facilities than those in the
279 community. The determined CoDs for children aged < 1 year (0.69) were higher than those
280 aged 1-4 years (0.28) despite their nearly no agreement when further stratified as stillbirths,

281 neonates, and infants. However, the agreement considerably increased when those groups
282 were combined (**Figure 3**).

283 Cause-specific mortality fractions determined by InterVA4 and DeCoDe differed in
284 important ways for some of the more common diseases (Table 3). In those surveillance sites
285 where the DeCoDe panels determined HIV as the underlying CoD for some deaths, the
286 InterVA4 model predicted considerably fewer HIV deaths, as demonstrated by CSMF. This
287 pattern is also seen in many sites for diagnoses such as malnutrition (Ethiopia, Kenya, Mali,
288 and Sierra Leone), neonatal sepsis (all sites except Mali), and birth asphyxia (all sites).
289 However, the InterVA4 predicted a substantially higher proportion of deaths caused in most
290 sites by prematurity (all sites, except Mali), malaria (Kenya, Mali, Mozambique, and Sierra
291 Leone), malnutrition (Mozambique and South Africa), diarrheal diseases (except in
292 Mozambique, which was lower), and meningitis (Kenya, Mozambique, Sierra Leone, and
293 South Africa) than did DeCoDe.

294 Some CoDs were exclusively assigned by the InterVA4 model rather than the
295 postmortem DeCoDe. For example, acute abdomen, renal failure, dengue fever, stroke, road
296 and other traffic accidents, accidental falls, and exposure to the force of nature were
297 exclusively assigned by the interVA4 model as the underlying CoDs and were not determined
298 by the experts using postmortem MITS. Conversely, unspecified external causes of death,
299 unspecified non-communicable diseases, congenital anomalies, digestive neoplasms, and
300 unspecified neoplasms were exclusively determined as CoDs by experts using postmortem
301 MITS. In addition, only the DeCoDe panels ascertained pulmonary tuberculosis as the CoD
302 in South Africa. At the same time, the InterVA4 model did not predict it. The InterVA4
303 model exclusively assigned epilepsy in Ethiopia, Kenya, Mali, and Sierra Leone, but both
304 InterVA4 and DeCoDe noted epilepsy as an underlying CoD in Mozambique and South
305 Africa.

306 The InterVA4 CSMF and DeCoDe CSMF ratios were calculated with a 95%
307 confidence interval using the CSMFs tabulated in Table 3 to show that many of these
308 differences did not occur by chance. The InterVA4 CoDs of fresh and macerated stillbirths
309 had the highest CSMF ratio, and unspecified neonatal CoD had the lowest CSMF ratio.
310

311 **Discussion**

312 This study compared the InterVA4 model with experts' determination of CoDs using
313 advanced diagnostics and postmortem MITS and showed poor InterVA4 agreement and
314 concordance in predicting the causes of death against DeCoDe among our <5s studied deaths.
315 The concordance suffered from its accuracy (< 0.8), although the precision was good (>0.8).

316 Several other studies compared the InterVA4 with Physicians-Certified Verbal
317 Autopsy (PCVA) and other standardized verbal autopsy diagnoses for public health
318 equivalence to test its functionality and costs [16, 47-50]. Others have also studied the
319 performance of InterVA4 with postmortem histologic findings. Knowing whether these tools
320 lead to similar conclusions—and if not, how results differ-- is important before relying on
321 verbal autopsy-generated information as the general country-wide source of CoD and for
322 planning and executing public health interventions [13]. This concept is particularly crucial in
323 a setting without widespread mortality registrations and in resource-constrained areas.

324 Across surveillance sites, there were considerable differences in the two systems'
325 concordance, as Ethiopia and Mozambique's LCC were good (>0.8) while the others were
326 poor. Our findings could be explained by quality differences in collecting the VA data and
327 the extent of CHAMPS's concurrent utilization of VA data with other clinical information to
328 assign the CoDs. The considerable agreement differences between children enrolled in health
329 institutions and the community also substantiate the argument, as death enrolled in the health
330 facility would have rich clinical information besides the VA compared to those cases enrolled
331 in the community. Furthermore, other studies have reported that the extent and way of VA
332 data collection determined how the InterVA4 assigned the respective CoDs [10, 16, 47, 51,
333 52]

334 However, we found the overall agreement in assigning the CoDs between the two
335 systems to be poor. This finding is unsurprising as several studies also found significant

336 differences between InterVA4 and PCVA or histologic findings [48, 53]. The concordance of
337 InterVA4 considerably decreases for stillbirths, neonates, and infants at individual and
338 population levels [33]. However, when they are combined, the level of agreement improves
339 significantly. Most importantly, more than a quarter of the overall sample were stillbirths,
340 where the InterVA4 is not designed to predict the causes of death. For example, most of the
341 diagnoses assigned by the InterVA4 for stillbirths were VAs-11.01 or VAs-11.02. These
342 assigned “macrated or fresh stillbirths” corresponded to the ICD-10 code of P95. In
343 addition, InterVA4 did not assign congenital anomalies arising during the prenatal period,
344 limiting its CoD equivalence compared to DeCoDe’s.

345 However, our findings did not agree with other studies that indicated an excellent
346 concordance between the assigned causes of death between the InterVA4 model and several
347 PCVA findings [47]. InterVA4 performed well in identifying malnutrition and certain
348 perinatal conditions as the underlying CoDs, similarly to the DeCoDe. For example, the ratio
349 of the proportion of malnutrition, birth asphyxia, and prematurity was closer to one or slightly
350 higher, meaning better equivalence in assigning those conditions..

351 Furthermore, the DeCoDe captures the overall mortality chain from underlying,
352 intermediate, and immediate causes of death, which is not done with InterVA4. In this study,
353 we could only compare the InterVA4 models’ most likely underlying causes of death to the
354 underlying causes attributed to DeCoDe. Comparing only the underlying CoDs may
355 potentially limit the overall correlation of causes of death between the two approaches, as
356 many deaths in live-born children occur after a complicated course of multiple causes [54].
357 For example, a neonate born prematurely could die of sepsis after admission to an intensive
358 care unit; in this case, DeCoDe would account for both causes. The InterVA4, however,
359 would mostly likely predict either of the causal chains, missing the overall causal chain.
360 These complete causal-chain scenarios identified by the DeCoDe panel would be based on

361 pieces of evidence from MITS and microbiological, clinical, and VA data. The DeCoDe
362 process does involve clinical judgment in some cases, as attributing causes of death from
363 multiple results can be complex, and clinical information, in particular, can be incomplete,
364 incorrect, or absent [55]. Nonetheless, errors should be few as the procedure is designed to
365 use the best possible set of information.

366 Another difference between the two methods is that the InterVA4 model mostly tends
367 to assign the stillbirths – either fresh or macerated—underlying CoDs in 80 % of the cases,
368 while the remaining CoDs designated were prematurity and intrauterine hypoxia; these CoDs,
369 that the InterVA4 mainly assigned for stillbirths were VA-11.01 and VA-11.02 corresponds
370 to the ICD code P95 – undetermined or unspecified causes of death. There were substantial
371 differences in assigning ICD code P95 between the two methods unrelated to chance, as the
372 InterVA4 assigned more than the DeCoDe. The differences could arise from the VA data
373 quality, the algorithm design, or the MITS's accuracy in determining the most likely causes of
374 death.

375 Similarly, prematurity was also more often assigned by InterVA4 than the DeCoDe
376 across sites, which could also be related to the level of certainty in determining the
377 pregnancy's gestational age, which the DeCoDe panel uses when assigning prenatal
378 mortality. Moreover, those babies born prematurely are most likely to have impending birth
379 asphyxia or respiratory distress, and the DeCoDe panel assigned more birth asphyxia than the
380 InterVA4 model. These differences point out the relationship and the complexities of the
381 causal chain that were responsible for the <5s deaths.

382 ***Conclusion***

383 Our findings point out that VA diagnosis alone, as generated by InterVA4, often
384 incorrectly predicts causes of death among <5s, using DeCoDe findings as the gold standard.
385 The InterVA4 model lacks precision in determining the underlying causes of death and

386 cannot predict some conditions like congenital anomalies. Future improvement of the
387 reliability and validity of VA data by strengthening the quality of data collection and
388 automatically assigning CoDs using robust and new technologies, such as artificial
389 intelligence, is recommended. Improving models to better predict causes of death, perhaps by
390 using information from deaths that also have information from postmortem diagnostic
391 assessments such as DeCoDe, would improve the usefulness of VA as a tool to inform health
392 policies [56, 57].

393 Overall, the role of the VA as a tool for diagnosing and tracking the progress of mortality data
394 among U5s is essential despite the noted shortcomings. Using the DeCoDe process that combines
395 Minimally invasive tissue sampling (MITS) and other techniques could provide data to help
396 improve CoDs determination. The data should subsequently be utilized to improve the CoD
397 determination algorithms of VA and its diagnostic ability.

398

399 **Acknowledgments**

400 We would like to thank all the community members from the CHAMPS sites and their local
401 authorities for their support. We thank all the CHAMPS teams on the sites, including their
402 legal representatives, for their support, the researchers, and the field staff involved in this
403 project. We also thank the Program office of the CHAMPS project at Emory University for
404 their continuous support and the local funders of each HDSS. Finally we would like to thank
405 Zachary Madewell and Beth Tippet Barr for providing their feedback in data handling and
406 drafting of the manuscript, respectively.

407 ***Author contributions:** NA, DB, and CY were involved in every stage of the study. All*
408 *co-authors, drafted, commented, edited, and approved submission.*

409 ***Ethics and consent***

410 Ethical clearances from the respective institutions and national ethical clearance bodies have
411 been secured for the HDSS and CHAMPS activities. The HDSS activities have standing
412 approval for continuing activities, including VA. The ethical review committee of the
413 respective Institutional Research Boards approved the study procedure at the sites. For all
414 participants, informed, voluntary, written, and signed consent was obtained from the
415 responsible person in the family (usually the head, the mother, or eligible family members).
416 Written informed consent was obtained from the parent/guardian of each participant under 18
417 years of age. We did not share data containing participants' identifiers with a third party to
418 maintain confidentiality.

419 **Disclaimer:** The findings and conclusions in this report are those of the author(s) and do not
420 necessarily represent the views of the U.S. Centers for Disease Control and Prevention.

421 **Funding statement:** CHAMPS is funded by the Bill and Melinda Gates Foundation (OPP1126780 to
422 CGW) which provided input into site selection decisions decision, methodology, and scope of
423 CHAMPS. The funders had no role in study design data collection and analysis, decision to publish,
424 or preparation of the manuscript. None of the authors receive a salary any of the funders.

425 **Declaration of interest statement:** No potential conflict of interest was reported by the author(s)

426 **Data availability statement:** The data supporting this study's findings are available on request

427 from the corresponding author, [Nega Assefa].

428 **ORCID:** Nega Assefa <https://orcid.org/0000-0003-0341-2329>

429 **Supporting information**

430 **S1 Figure:** Flowchart of deaths included in the analysis

431 **S2 Figure:** Concordance of cause-specific mortality fractions of the underlying causes of death
432 between InterVA4 and DeCoDe for CHAMPS surveillance sites

433 **S3 Figure:** Concordance of cause-specific mortality fractions of the underlying causes of death
434 between InterVA4 and DeCoDe for CHAMPS surveillance sites by age group.

435

436 **References**

- 437 1. Fottrell E, Byass P. Verbal autopsy: methods in transition. *Epidemiol Rev.*
438 2010;32(1):38-55. Epub 20100304. doi: 10.1093/epirev/mxq003. PubMed PMID: 20203105.
- 439 2. Wallace AS, Ryman TK, Dietz V. Overview of global, regional, and national routine
440 vaccination coverage trends and growth patterns from 1980 to 2009: implications for vaccine-
441 preventable disease eradication and elimination initiatives. *J Infect Dis.* 2014;210 Suppl 1(0
442 1):S514-22. doi: 10.1093/infdis/jiu108. PubMed PMID: 25316875; PubMed Central PMCID:
443 PMCPMC4663673.
- 444 3. Alkema L, You D. Child mortality estimation: a comparison of UN IGME and IHME
445 estimates of levels and trends in under-five mortality rates and deaths. *PLoS Med.*
446 2012;9(8):e1001288. doi: 10.1371/journal.pmed.1001288. PubMed PMID: 22952434;
447 PubMed Central PMCID: PMCPMC3429386.
- 448 4. Mathers CD, Fat DM, Inoue M, Rao C, Lopez AD. Counting the dead and what they
449 died from: an assessment of the global status of cause of death data. *Bull World Health*
450 *Organ.* 2005;83(3):171-7. Epub 20050316. doi: /S0042-96862005000300009. PubMed
451 PMID: 15798840; PubMed Central PMCID: PMCPMC2624200.
- 452 5. Lohela TJ, Nesbitt RC, Pekkanen J, Gabrysch S. Comparing socioeconomic
453 inequalities between early neonatal mortality and facility delivery: Cross-sectional data from
454 72 low- and middle-income countries. *Sci Rep.* 2019;9(1):9786. Epub 20190705. doi:
455 10.1038/s41598-019-45148-5. PubMed PMID: 31278283; PubMed Central PMCID:
456 PMCPMC6611781.
- 457 6. Liu L, Li Q, Lee RA, Friberg IK, Perin J, Walker N, et al. Trends in causes of death
458 among children under 5 in Bangladesh, 1993-2004: an exercise applying a standardized
459 computer algorithm to assign causes of death using verbal autopsy data. *Popul Health Metr.*
460 2011;9(1):43. doi: 10.1186/1478-7954-9-43. PubMed PMID: 21819600; PubMed Central
461 PMCID: PMCPMC3160936.
- 462 7. Lulu K, Berhane Y. The use of simplified verbal autopsy in identifying causes of
463 adult death in a predominantly rural population in Ethiopia. *BMC Public Health.* 2005;5:58.

464 Epub 20050603. doi: 10.1186/1471-2458-5-58. PubMed PMID: 15935096; PubMed Central
465 PMCID: PMCPMC1164421.

466 8. Setel PW, Whiting DR, Hemed Y, Chandramohan D, Wolfson LJ, Alberti KG, et al.
467 Validity of verbal autopsy procedures for determining cause of death in Tanzania. *Trop Med*
468 *Int Health*. 2006;11(5):681-96. doi: 10.1111/j.1365-3156.2006.01603.x. PubMed PMID:
469 16640621.

470 9. Abbas SM, Alam AY, Majid A. To determine the probable causes of death in an
471 urban slum community of Pakistan among adults 18 years and above by verbal autopsy.
472 *JPMMA The Journal of the Pakistan Medical Association*. 2011;61(3):235-8. Epub 2011/04/07.
473 PubMed PMID: 21465935.

474 10. Morris SK, Bassani DG, Kumar R, Awasthi S, Paul VK, Jha P. Factors associated
475 with physician agreement on verbal autopsy of over 27000 childhood deaths in India. *PLoS*
476 *One*. 2010;5(3):e9583. Epub 20100308. doi: 10.1371/journal.pone.0009583. PubMed PMID:
477 20221398; PubMed Central PMCID: PMCPMC2833201.

478 11. Chandramohan D, Fottrell E, Leitao J, Nichols E, Clark SJ, Alsokhn C, et al.
479 Estimating causes of death where there is no medical certification: evolution and state of the
480 art of verbal autopsy. *Global health action*. 2021;14(sup1):1982486. doi:
481 10.1080/16549716.2021.1982486. PubMed PMID: 35377290; PubMed Central PMCID:
482 PMCPMC8986278.

483 12. Quigley MA, Armstrong Schellenberg JR, Snow RW. Algorithms for verbal
484 autopsies: a validation study in Kenyan children. *Bull World Health Organ*. 1996;74(2):147-
485 54. Epub 1996/01/01. PubMed PMID: 8706229; PubMed Central PMCID:
486 PMCPMC2486900.

487 13. Foreman KJ, Lozano R, Lopez AD, Murray CJ. Modeling causes of death: an
488 integrated approach using CODEm. *Popul Health Metr*. 2012;10:1. Epub 20120106. doi:
489 10.1186/1478-7954-10-1. PubMed PMID: 22226226; PubMed Central PMCID:
490 PMCPMC3315398.

- 491 14. Allotey PA, Reidpath DD, Evans NC, Devarajan N, Rajagobal K, Bachok R, et al.
492 Let's talk about death: data collection for verbal autopsies in a demographic and health
493 surveillance site in Malaysia. *Global health action*. 2015;8(1):28219. Epub 20150701. doi:
494 10.3402/gha.v8.28219. PubMed PMID: 26140728; PubMed Central PMCID:
495 PMCPMC4490796.
- 496 15. Nabukalu D, Ntaro M, Seviiri M, Reyes R, Wiens M, Sundararajan R, et al.
497 Community health workers trained to conduct verbal autopsies provide better mortality
498 measures than existing surveillance: Results from a cross-sectional study in rural western
499 Uganda. *PLoS One*. 2019;14(2):e0211482. Epub 20190213. doi:
500 10.1371/journal.pone.0211482. PubMed PMID: 30759139; PubMed Central PMCID:
501 PMCPMC6373919.
- 502 16. Uneke CJ, Uro-Chukwu HC, Chukwu OE. Validation of verbal autopsy methods for
503 assessment of child mortality in sub-Saharan Africa and the policy implication: a rapid
504 review. *The Pan African medical journal*. 2019;33:318. Epub 20190822. doi:
505 10.11604/pamj.2019.33.318.16405. PubMed PMID: 31692720; PubMed Central PMCID:
506 PMCPMC6815483.
- 507 17. Mahesh BPK, Hart JD, Acharya A, Chowdhury HR, Joshi R, Adair T, et al.
508 Validation studies of verbal autopsy methods: a systematic review. *BMC Public Health*.
509 2022;22(1):2215. Epub 20221129. doi: 10.1186/s12889-022-14628-1. PubMed PMID:
510 36447199; PubMed Central PMCID: PMCPMC9706899.
- 511 18. Garenne M. Validations of verbal autopsies - comment on 'automated versus
512 physician assignment of cause of death for verbal autopsies: randomized trial of 9374 deaths
513 in 117 villages in India'. *BMC Med*. 2019;17(1):131. doi: 10.1186/s12916-019-1371-0.
514 PubMed PMID: 31307456; PubMed Central PMCID: PMCPMC6633696.
- 515 19. Kuniyama T, Li ZR, Clark SJ, McCormick TH. Bayesian Factor Models for
516 Probabilistic Cause of Death Assessment with Verbal Autopsies. *Ann Appl Stat*.
517 2020;14(1):241-56. Epub 04/16. doi: 10.1214/19-aos1253. PubMed PMID: 33520049;
518 PubMed Central PMCID: PMCPMC7845920.

- 519 20. Murray CJ, Lozano R, Flaxman AD, Serina P, Phillips D, Stewart A, et al. Using
520 verbal autopsy to measure causes of death: the comparative performance of existing methods.
521 BMC Med. 2014;12(1):5. doi: 10.1186/1741-7015-12-5. PubMed PMID: 24405531; PubMed
522 Central PMCID: PMCPMC3891983.
- 523 21. Patterson JK, Aziz A, Bauserman MS, McClure EM, Goldenberg RL, Bose CL.
524 Challenges in classification and assignment of causes of stillbirths in low- and lower middle-
525 income countries. *Semin Perinatol.* 2019;43(5):308-14. Epub 20190316. doi:
526 10.1053/j.semperi.2019.03.021. PubMed PMID: 30981473; PubMed Central PMCID:
527 PMCPMC7894980.
- 528 22. Halim A, Aminu M, Dewez JE, Biswas A, Rahman A, van den Broek N. Stillbirth
529 surveillance and review in rural districts in Bangladesh. *BMC Pregnancy Childbirth.*
530 2018;18(1):224. Epub 20180613. doi: 10.1186/s12884-018-1866-2. PubMed PMID:
531 29914393; PubMed Central PMCID: PMCPMC6004696.
- 532 23. Paganelli CR, Goco NJ, McClure EM, Banke KK, Blau DM, Breiman RF, et al. The
533 evolution of minimally invasive tissue sampling in postmortem examination: a narrative
534 review. *Global health action.* 2020;13(1):1792682. doi: 10.1080/16549716.2020.1792682.
535 PubMed PMID: 32713325; PubMed Central PMCID: PMCPMC7480574.
- 536 24. Menendez C, Castillo P, Martinez MJ, Jordao D, Lovane L, Ismail MR, et al. Validity
537 of a minimally invasive autopsy for cause of death determination in stillborn babies and
538 neonates in Mozambique: An observational study. *PLoS Med.* 2017;14(6):e1002318. doi:
539 10.1371/journal.pmed.1002318. PubMed PMID: 28632735; PubMed Central PMCID:
540 PMCPMC5478138.
- 541 25. Chawana R, Baillie V, Izu A, Solomon F, Bassat Q, Blau DM, et al. Potential of
542 Minimally Invasive Tissue Sampling for Attributing Specific Causes of Childhood Deaths in
543 South Africa: A Pilot, Epidemiological Study. *Clinical infectious diseases : an official
544 publication of the Infectious Diseases Society of America.* 2019;69(Suppl 4):S361-S73. doi:
545 10.1093/cid/ciz550. PubMed PMID: 31598659; PubMed Central PMCID:
546 PMCPMC6785686.

- 547 26. Blevins J, O'Mara Sage E, Kone A, Maixenchs M, Raghunathan PL, Guilaze RA, et
548 al. Using Participatory Workshops to Assess Alignment or Tension in the Community for
549 Minimally Invasive Tissue Sampling Prior to Start of Child Mortality Surveillance: Lessons
550 From 5 Sites Across the CHAMPS Network. *Clinical infectious diseases : an official*
551 *publication of the Infectious Diseases Society of America.* 2019;69(Suppl 4):S280-S90. doi:
552 10.1093/cid/ciz563. PubMed PMID: 31598665; PubMed Central PMCID:
553 PMC6785692.
- 554 27. Flaxman AD, Stewart A, Joseph JC, Alam N, Alam SS, Chowdhury H, et al.
555 Collecting verbal autopsies: improving and streamlining data collection processes using
556 electronic tablets. *Popul Health Metr.* 2018;16(1):3. Epub 20180201. doi: 10.1186/s12963-
557 018-0161-9. PubMed PMID: 29391038; PubMed Central PMCID: PMC6785692.
- 558 28. Raghunathan PL, Madhi SA, Breiman RF. Illuminating Child Mortality: Discovering
559 Why Children Die. *Clinical infectious diseases : an official publication of the Infectious*
560 *Diseases Society of America.* 2019;69(Suppl 4):S257-S9. doi: 10.1093/cid/ciz562. PubMed
561 PMID: 31598658; PubMed Central PMCID: PMC6785666.
- 562 29. Cunningham SA, Shaikh NI, Nacolo A, Raghunathan PL, Kotloff K, Naser AM, et
563 al. Health and Demographic Surveillance Systems Within the Child Health and Mortality
564 Prevention Surveillance Network. *Clinical infectious diseases : an official publication of the*
565 *Infectious Diseases Society of America.* 2019;69(Suppl 4):S274-S9. doi: 10.1093/cid/ciz609.
566 PubMed PMID: 31598663; PubMed Central PMCID: PMC6785673.
- 567 30. Dowell SF, Zaidi A, Heaton P. Why Child Health and Mortality Prevention
568 Surveillance? *Clinical infectious diseases : an official publication of the Infectious Diseases*
569 *Society of America.* 2019;69(Suppl 4):S260-S1. doi: 10.1093/cid/ciz542. PubMed PMID:
570 31598662; PubMed Central PMCID: PMC6785698.
- 571 31. Salzberg NT, Sivalogan K, Bassat Q, Taylor AW, Adedini S, El Arifeen S, et al.
572 Mortality Surveillance Methods to Identify and Characterize Deaths in Child Health and
573 Mortality Prevention Surveillance Network Sites. *Clinical infectious diseases : an official*
574 *publication of the Infectious Diseases Society of America.* 2019;69(Suppl 4):S262-S73. doi:

- 575 10.1093/cid/ciz599. PubMed PMID: 31598664; PubMed Central PMCID:
576 PMC6785672.
- 577 32. World Health Organization. Verbal autopsy standards : ascertaining and attributing
578 cause of death. Geneva: World Health Organization; 2007.
- 579 33. Lozano R, Freeman MK, James SL, Campbell B, Lopez AD, Flaxman AD, et al.
580 Performance of InterVA for assigning causes of death to verbal autopsies: multisite validation
581 study using clinical diagnostic gold standards. *Popul Health Metr.* 2011;9(1):50. Epub
582 20110805. doi: 10.1186/1478-7954-9-50. PubMed PMID: 21819580; PubMed Central
583 PMCID: PMC3160943.
- 584 34. Richard Li Z, Thomas J, Choi E, H. McCormick T, J Clark S. The openVA Toolkit
585 for Verbal Autopsies. *The R Journal.* 2023;14(4):316-34. doi: 10.32614/rj-2023-020.
- 586 35. World Health Organization. The WHO application of ICD-10 to deaths during the
587 perinatal period: ICD-PM. Geneva: World Health Organization; 2016.
- 588 36. World Health Organization. Verbal autopsy standards: the 2016 WHO verbal autopsy
589 instrument. Geneva, Switzerland: WHO; 2016.
- 590 37. World Health Organization. International Statistical Classification of Diseases and
591 related health problems: Alphabetical index: World Health Organization; 2004.
- 592 38. CHAMPS. CHAMPS Diagnosis standards [July 7, 2023]. Available from:
593 <https://champshealth.org/resources/champs-diagnosis-standards/>.
- 594 39. World Health Organization. Verbal autopsy standards: The 2022 WHO verbal
595 autopsy instrument. Geneva: World Health Organization; 2017.
- 596 40. Pathirana J, Munoz FM, Abbing-Karahagopian V, Bhat N, Harris T, Kapoor A, et al.
597 Neonatal death: Case definition & guidelines for data collection, analysis, and presentation of
598 immunization safety data. *Vaccine.* 2016;34(49):6027-37. Epub 20160719. doi:
599 10.1016/j.vaccine.2016.03.040. PubMed PMID: 27449077; PubMed Central PMCID:
600 PMC5139812.

- 601 41. Weldearegawi B, Melaku YA, Abera SF, Ashebir Y, Haile F, Mulugeta A, et al.
602 Infant mortality and causes of infant deaths in rural Ethiopia: a population-based cohort of
603 3684 births. *BMC Public Health*. 2015;15(1):770. Epub 20150811. doi: 10.1186/s12889-015-
604 2090-x. PubMed PMID: 26260495; PubMed Central PMCID: PMC4531534.
- 605 42. Barnhart HX, Haber M, Song J. Overall concordance correlation coefficient for
606 evaluating agreement among multiple observers. *Biometrics*. 2002;58(4):1020-7. doi:
607 10.1111/j.0006-341x.2002.01020.x. PubMed PMID: 12495158.
- 608 43. Cox N, Steichen T. CONCORD: Stata module for concordance correlation. *Statistical*
609 *Software Components*: Boston College Department of Economics; 2007.
- 610 44. Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments
611 used in research. *Am J Health Syst Pharm*. 2008;65(23):2276-84. doi: 10.2146/ajhp070364.
612 PubMed PMID: 19020196.
- 613 45. Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med*. 2018;18(3):91-
614 3. Epub 20180807. doi: 10.1016/j.tjem.2018.08.001. PubMed PMID: 30191186; PubMed
615 Central PMCID: PMC6107969.
- 616 46. Fagerland MW, Lydersen S, Laake P. Recommended confidence intervals for two
617 independent binomial proportions. *Stat Methods Med Res*. 2015;24(2):224-54. Epub
618 20111013. doi: 10.1177/0962280211415469. PubMed PMID: 21996567.
- 619 47. Byass P, Herbst K, Fottrell E, Ali MM, Odhiambo F, Amek N, et al. Comparing
620 verbal autopsy cause of death findings as determined by physician coding and probabilistic
621 modelling: a public health analysis of 54 000 deaths in Africa and Asia. *J Glob Health*.
622 2015;5(1):010402. doi: 10.7189/jogh.05.010402. PubMed PMID: 25734004; PubMed
623 Central PMCID: PMC4337147.
- 624 48. Vergnano S, Fottrell E, Osrin D, Kazembe PN, Mwansambo C, Manandhar DS, et al.
625 Adaptation of a probabilistic method (InterVA) of verbal autopsy to improve the
626 interpretation of cause of stillbirth and neonatal death in Malawi, Nepal, and Zimbabwe.

627 Popul Health Metr. 2011;9:48. Epub 20110805. doi: 10.1186/1478-7954-9-48. PubMed
628 PMID: 21819599; PubMed Central PMCID: PMCPMC3160941.

629 49. Weldearegawi B, Melaku YA, Dinant GJ, Spigt M. How much do the physician
630 review and InterVA model agree in determining causes of death? A comparative analysis of
631 deaths in rural Ethiopia. BMC Public Health. 2015;15(1):669. Epub 20150715. doi:
632 10.1186/s12889-015-2032-7. PubMed PMID: 26173990; PubMed Central PMCID:
633 PMCPMC4503295.

634 50. Tadesse S. Validating the InterVA model to estimate the burden of mortality from
635 verbal autopsy data: a population-based cross-sectional study. PLoS One. 2013;8(9):e73463.
636 Epub 20130913. doi: 10.1371/journal.pone.0073463. PubMed PMID: 24058474; PubMed
637 Central PMCID: PMCPMC3772846.

638 51. Redelings MD, Sorvillo F, Simon P. A Comparison of Underlying Cause and
639 Multiple Causes of Death: US Vital Statistics, 2000–2001. Epidemiology. 2006;17(1).

640 52. Ronsmans C, Vanneste AM, Chakraborty J, Van Ginneken J. A comparison of three
641 verbal autopsy methods to ascertain levels and causes of maternal deaths in Matlab,
642 Bangladesh. Int J Epidemiol. 1998;27(4):660-6. doi: 10.1093/ije/27.4.660. PubMed PMID:
643 9758122.

644 53. Roulson J, Benbow EW, Hasleton PS. Discrepancies between clinical and autopsy
645 diagnosis and the value of post mortem histology; a meta-analysis and review.
646 Histopathology. 2005;47(6):551-9. doi: 10.1111/j.1365-2559.2005.02243.x. PubMed PMID:
647 16324191.

648 54. Menéndez C, Quintó L, Castillo P, Carrilho C, Ismail MR, Lorenzoni C, et al.
649 Limitations to current methods to estimate cause of death: a validation study of a verbal
650 autopsy model. Gates Open Res. 2020;4:55. Epub 20210505. doi:
651 10.12688/gatesopenres.13132.3. PubMed PMID: 33145479; PubMed Central PMCID:
652 PMCPMC7590499.

- 653 55. Byass P. Whither verbal autopsy? *Popul Health Metr.* 2011;9(1):23. Epub 20110801.
654 doi: 10.1186/1478-7954-9-23. PubMed PMID: 21806833; PubMed Central PMCID:
655 PMCPMC3160916.
- 656 56. Fiksel J, Gilbert B, Wilson E, Kalter H, Kante A, Akum A, et al. Correcting for
657 Verbal Autopsy Misclassification Bias in Cause-Specific Mortality Estimates. *Am J Trop*
658 *Med Hyg.* 2023;108(5_Suppl):66-77. Epub 20230410. doi: 10.4269/ajtmh.22-0318. PubMed
659 PMID: 37037438; PubMed Central PMCID: PMCPMC10160858.
- 660 57. Gilbert B, Fiksel J, Wilson E, Kalter H, Kante A, Akum A, et al. Multi-Cause
661 Calibration of Verbal Autopsy-Based Cause-Specific Mortality Estimates of Children and
662 Neonates in Mozambique. *Am J Trop Med Hyg.* 2023;108(5_Suppl):78-89. Epub 20230410.
663 doi: 10.4269/ajtmh.22-0319. PubMed PMID: 37037430; PubMed Central PMCID:
664 PMCPMC10160855.
665

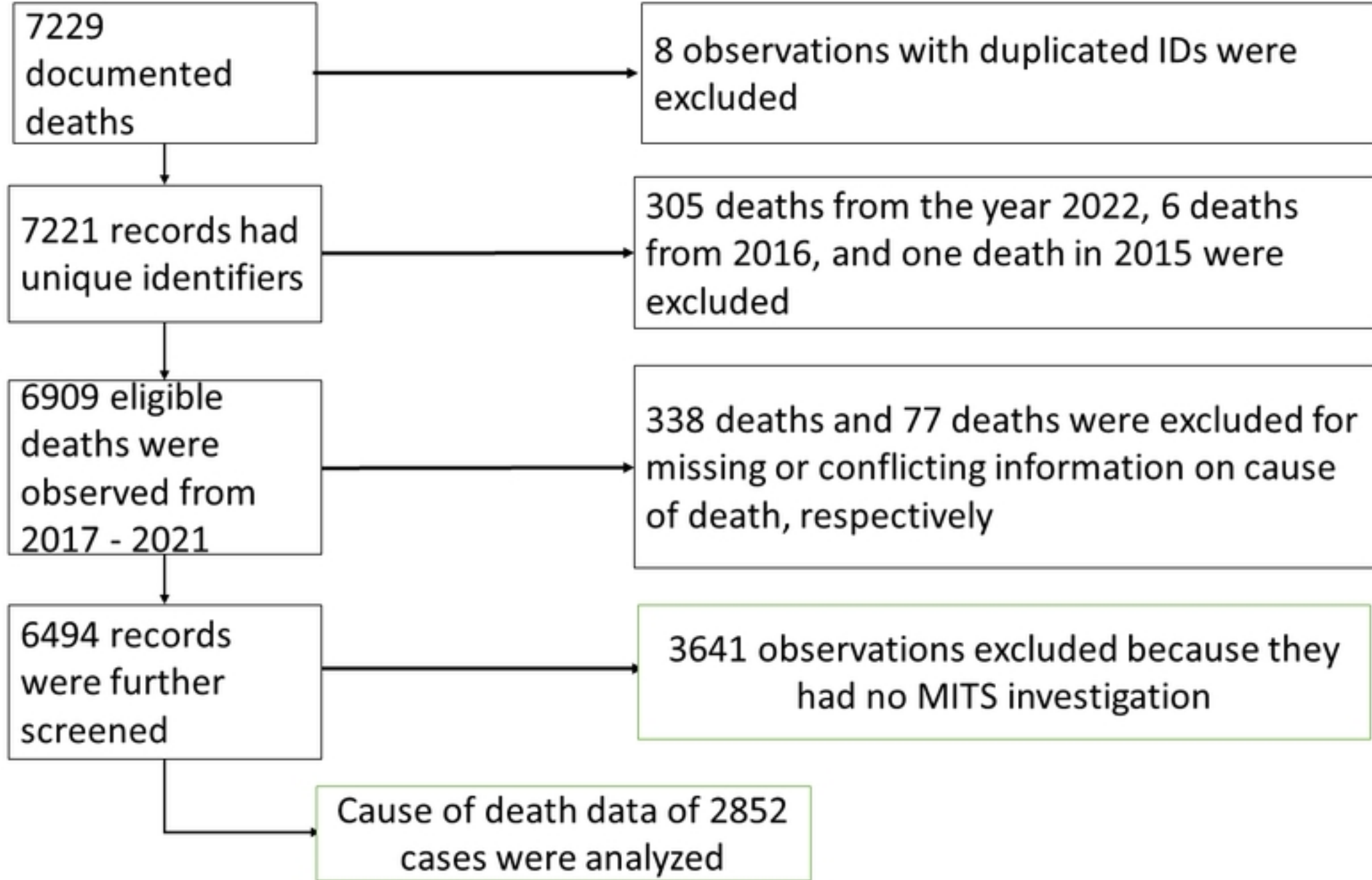


Figure 1

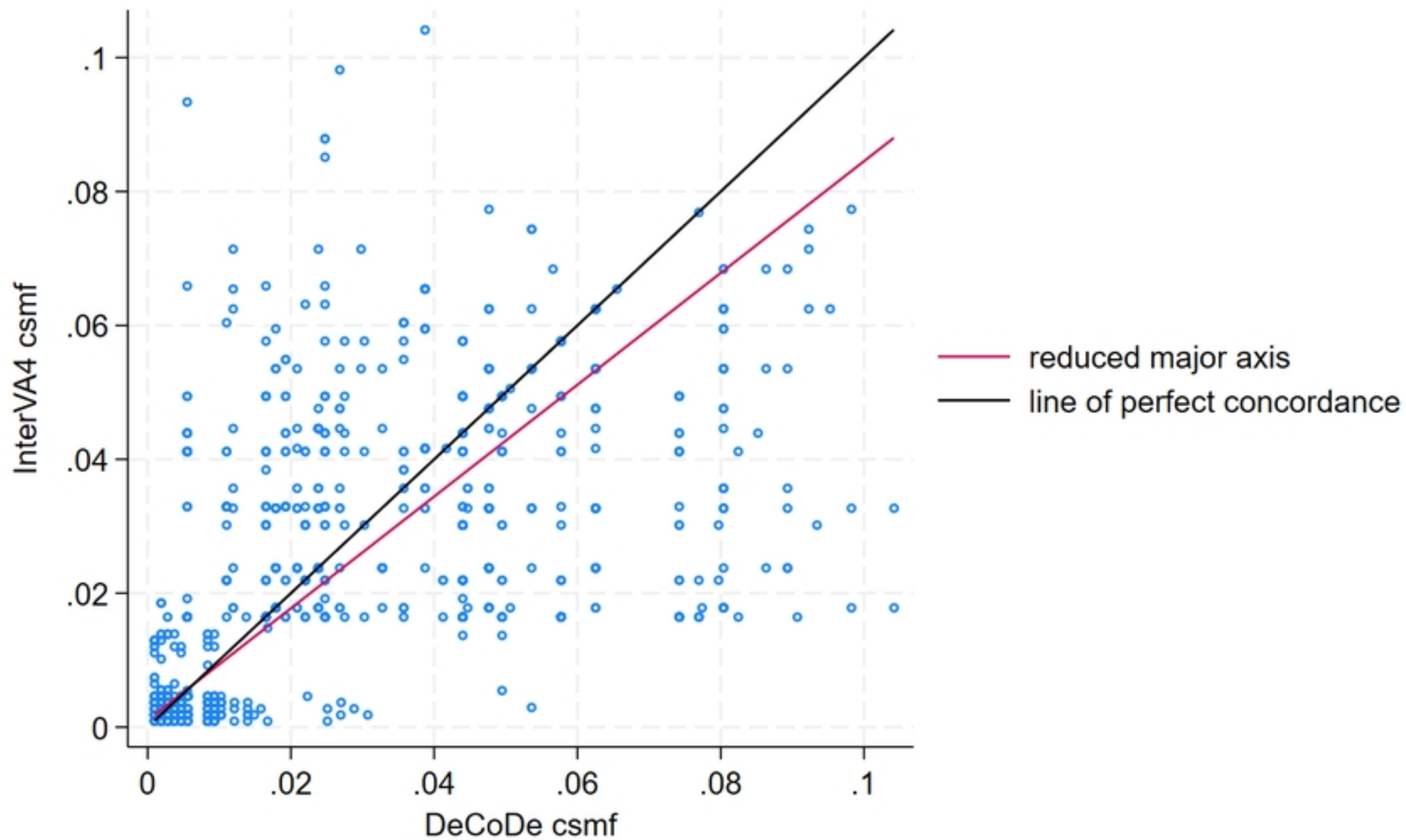


Figure 2

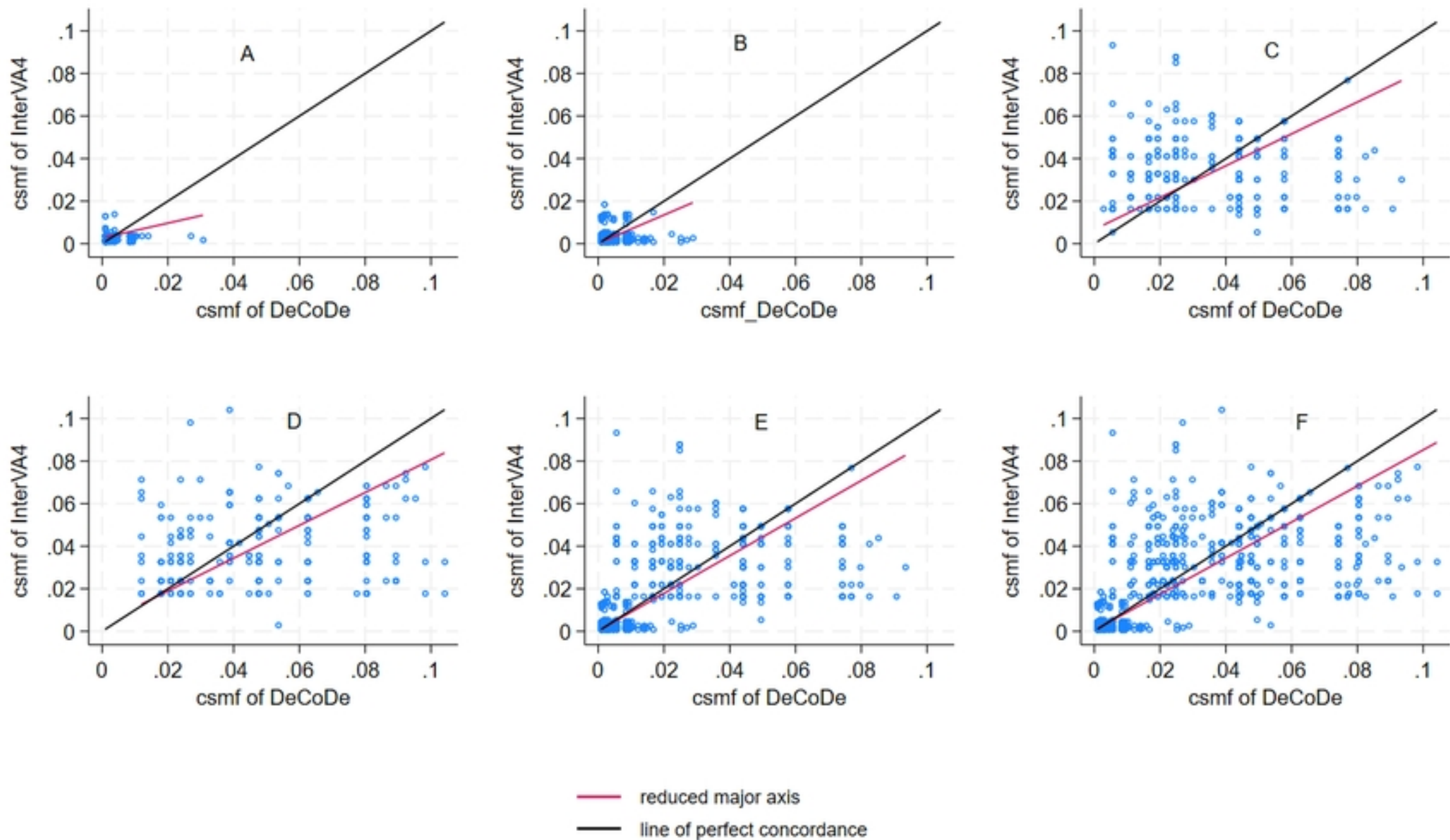


Figure 3