

1 **Towards estimating true cholera burden: a systematic review and meta-**
2 **analysis of *Vibrio cholerae* positivity**

3

4 **Running title:** Estimating *Vibrio cholerae* positivity

5

6 Kirsten E. Wiens^{1,2}, Hanmeng Xu¹, Kaiyue Zou¹, John Mwaba^{3,4}, Justin Lessler^{1,5,6}, Espoir B.

7 Malembaka^{1,7}, Maya N. Demby¹, Godfrey Bwire⁸, Firdausi Qadri⁹, Elizabeth C. Lee¹, Andrew S.

8 Azman¹

9

10 ¹ Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Johns

11 Hopkins University, Baltimore, USA

12 ² Department of Epidemiology and Biostatistics, College of Public Health, Temple University,

13 Philadelphia, USA

14 ³ Centre for Infectious Disease Research in Zambia (CIDRZ), Lusaka, Zambia

15 ⁴ Department of Biomedical Sciences, School of Health Sciences, University of Zambia, Lusaka,

16 Zambia

17 ⁵ Department of Epidemiology, Gillings School of Global Public Health, University of North

18 Carolina at Chapel Hill, Chapel Hill, USA

19 ⁶ Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, USA

20 ⁷ Center for Tropical Diseases and Global Health (CTDGH), Université Catholique de Bukavu,

21 Bukavu, Democratic Republic of the Congo

22 ⁸ Division of Public Health Emergency Preparedness and Response, Ministry of Health,

23 Kampala, Uganda

24 ⁹ Infectious Diseases Division, International Centre for Diarrhoeal Disease Research

25 Bangladesh (icddr,b), Dhaka, Bangladesh

26 **Corresponding author**

27 Andrew S. Azman, PhD

28 Associate Scientist

29 Department of Epidemiology

30 Johns Hopkins Bloomberg School of Public Health

31 615 N. Wolfe Street

32 Baltimore, MD 21205

33 azman@jhu.edu

34

35 **Alternative corresponding author**

36 Kirsten E. Wiens, PhD

37 Assistant Professor

38 Department of Epidemiology and Biostatistics

39 Temple University College of Public Health

40 1301 Cecil B. Moore Avenue, Ritter Annex 911

41 Philadelphia, PA 19122

42 kirsten.wiens@temple.edu

43

44 **Key points**

45 By pooling results from 113 studies and adjusting for diagnostic tests and study methods, we find

46 that approximately half of suspected cholera cases represent true *Vibrio cholerae* infections, though

47 this fraction varies widely across epidemiological settings.

48

49 **Keywords**

50 cholera, *Vibrio cholerae*, diagnostic test, suspected cholera, confirmed cholera

51

52 **Abstract**

53 Background: Cholera surveillance relies on clinical diagnosis of acute watery diarrhea.

54 Suspected cholera case definitions have high sensitivity but low specificity, challenging our
55 ability to characterize cholera burden and epidemiology. Our objective was to estimate the
56 proportion of clinically suspected cholera that are true *Vibrio cholerae* infections and identify
57 factors that explain variation in positivity.

58
59 Methods: We conducted a systematic review of studies from 2000-2021 that tested ≥ 10
60 suspected cholera cases for *V. cholerae* O1/O139 using culture, PCR and/or a rapid diagnostic
61 test. We estimated diagnostic test sensitivity and specificity using a latent class meta-analysis.
62 We estimated positivity using a random-effects meta-analysis, adjusting for test performance
63 and study methodology.

64
65 Results: We included 113 studies from 28 countries. *V. cholerae* positivity was lower in studies
66 with representative sampling and lower minimum ages in suspected case definitions. After
67 adjusting for sampling methods, case definitions, and tests, on average half (49%, 95% Credible
68 Interval: 43%-54%) of suspected cases represented true *V. cholerae* infections, although
69 variation across studies was high. Odds of a suspected case having a true infection were 1.64
70 (95% Credible Interval: 1.06-2.52) times higher when surveillance was initiated in response to
71 an outbreak than in non-outbreak settings.

72
73 Conclusions: Burden estimates based on suspected cases alone may overestimate the
74 incidence of medically attended cholera about twofold. However, accounting for cases missed
75 by traditional clinical surveillance is key to unbiased overall cholera burden estimates. Given
76 variability between settings, assumptions about positivity, which are necessary without
77 exhaustive testing, should be based on local data.

78

79 Introduction

80 Current estimates of cholera burden rely on clinical diagnosis of individuals with acute
81 watery diarrhea (i.e., suspected cholera cases) [1,2]. It is unclear how many *Vibrio cholerae*
82 O1/O139 (serogroups that cause current epidemics) infections get missed due to mild
83 symptoms and other barriers to care-seeking or how many get overcounted due to non-specific
84 suspected case definitions. In Bangladesh, previous studies estimated that asymptomatic and
85 unreported infections account for at least half of *V. cholerae* infections [3–5]. Meanwhile, the
86 proportion of suspected cholera cases that represent laboratory-confirmed infections varies
87 widely between studies, from 6% of those tested during routine surveillance in Bangladesh [6]
88 to 72% of those tested during the initial phase of the 2017 outbreak in Yemen [7].

89 This wide variation in positivity may be caused by differences between sites in *V.*
90 *cholerae* epidemiology [8], epidemiology of non-cholera diseases causing the same clinical
91 symptoms [9–12], and variations in diagnostic tests and case definitions [13–15]. Typical
92 suspected cholera case definitions have been shown to have high sensitivity but low specificity
93 [14] and can vary by location across seasons [13]. Culture-based methods or PCR are the gold
94 standards to confirm cholera in clinical samples and generally have high specificity. Lateral flow
95 rapid diagnostic tests (RDT) may also be used and can be as sensitive as PCR [16]. Although
96 recommended by the Global Task Force on Cholera Control (GTFCC) [17], systematic
97 microbiological confirmation in surveillance is not always implemented, particularly during
98 outbreaks when resources are limited [8].

99 Understanding *V. cholerae* positivity among clinical cases could provide insights needed
100 to improve laboratory testing strategies and allow for better estimates of cholera burden and
101 risk, which are often used to allocate cholera resources, including oral cholera vaccines. We
102 sought to address this knowledge gap by modeling the relationship between clinically suspected
103 and laboratory confirmed cholera. Specifically, we aimed to estimate the proportion of

104 suspected cholera cases that represent true *V. cholerae* O1/O139 infections and identify factors
105 that explain variability in positivity across sites.

106

107 **Methods**

108 Terminology

109 We focused on *V. cholerae* O1 and O139 because these are the serogroups that are
110 responsible for the current 7th pandemic and the only ones known to lead to large outbreaks in
111 humans [18]. These are also the serogroups that are targeted by each of the commonly used *V.*
112 *cholerae* diagnostic tests (culture, PCR, RDT). Throughout this manuscript, we refer to the
113 proportion of suspected cholera cases that represent true *V. cholerae* O1/O139 infections as “*V.*
114 *cholerae* positivity” or “cholera positivity”. In addition, since the available data did not allow us to
115 evaluate performance of multiple RDTs, we refer to RDT as any rapid diagnostic test for *V.*
116 *cholerae* O1/O139 and do not distinguishing between different RDT manufacturers or whether
117 the RDT is enriched/direct swab RDT or stool RDT.

118

119 Systematic review

120 We searched PubMed, Embase, Scopus, Google Scholar and *medRxiv* on October 16,
121 2021 using search provided in the Supplementary Methods. We included studies that: 1)
122 collected human samples, 2) reported the number of suspected and confirmed cholera cases in
123 the sampling frame, 3) used culture, PCR, and/or RDT to test suspected cases for cholera, and
124 4) had at least one suspected case sample collected on or after January 1, 2000. We excluded
125 studies that: 1) used a case definition not specific for suspected cholera (i.e., we accepted non-
126 bloody watery diarrhea, acute watery diarrhea, or simply suspected cholera but not diarrhea,
127 acute diarrhea, or acute gastroenteritis), 2) sampled only special populations (i.e., people living
128 with HIV or cancer), 3) selected suspected cases based on epidemiological link to other cases

129 or environmental sources, 4) tested fewer than 10 suspected cases, 5) were reported in
130 languages other than English, French, Spanish, and Chinese.

131 Titles, abstracts, and full texts were uploaded to Covidence, a web-based screening tool
132 (<https://www.covidence.org/>), and were assessed independently by two of the reviewers (ASA,
133 ECL, HX, KEW, KZ, MND) for inclusion. Conflicts were resolved either by a third reviewer or
134 through consensus/discussion. Data were extracted from included studies in a shared
135 spreadsheet (Supplementary Data 1). The key extracted items included study timeframe and
136 location, surveillance type (routine, outbreak, post-vaccination, or hybrid), case definition of
137 suspected cholera (including age constraint and whether dehydrated or hospitalized, if
138 provided), test method(s), sampling strategy for the test (all suspected cases, systematic or
139 random sampling, convenience sampling, or unreported), number of tested and confirmed
140 suspected cases, among other sample characteristics, if included. If the surveillance contained
141 multiple timeframes, tested samples with multiple tests, or reported stratified results, we
142 extracted the data separately into different rows in the spreadsheet.

143 To identify overlapping samples, we manually reviewed all studies with overlapping
144 timeframes by country. We excluded studies that had shorter timeframes, fewer suspected
145 cases tested, less representative sampling methods, fewer confirmation tests, or reported
146 positive results by two tests but did not disaggregate. Within studies, when suspected and
147 confirmed cases were stratified multiple ways, we included the stratification by surveillance type
148 if available, followed by age, antibiotic use, dehydration status, year, geography, or sex, in that
149 order. When studies used multiple RDTs, we included results for Crystal VC (Arkray Healthcare
150 Pvt. Ltd, Gujarat, India) and direct rapid tests (as opposed to rapid tests performed after an
151 enrichment step) because these were the most common.

152

153 Data analysis

154 *Estimating sensitivity and specificity of cholera confirmation tests*

155 We constructed a latent-class model to assess sensitivity and specificity of culture, PCR,
156 and RDT, assuming none had perfect performance. We fit a hierarchical conditional
157 dependence model, similar to that proposed by Wang *et al.*, which takes into account potential
158 pairwise dependence between the tests that could occur if the tests have reduced performance
159 for similar reasons [19]. We performed inference in a Bayesian framework using Just Another
160 Gibbs Sampler (JAGS) through the *rjags* package in R [20,21]. We pooled estimates across four
161 published studies that reported cholera confirmation results for all three test methods [16,22–
162 24]. We used uninformative prior distributions on sensitivity and specificity of each test with a
163 lower bound set based on plausible values from the literature (Table S1). We ran 4 chains of
164 100,000 iterations and assessed convergence using the Gelman-Rubin R-hat statistic.

165

166 *Estimating V. cholerae positivity and sources of heterogeneity*

167 We pooled estimates of *V. cholerae* positivity across all studies using a meta-analysis
168 with hierarchical modeling, adjusting for sensitivity and specificity of the diagnostic tests as well
169 as study methodology and setting (see Supplementary Methods for the full statistical model).
170 We used the same statistical model to examine factors associated with variation in positivity.

171 To examine between- versus within-study heterogeneity, we calculated the I^2 statistic
172 [25] with

$$173 \quad I^2 = \frac{\tau^2}{\tau^2 + v}$$

174 where τ^2 was between-study heterogeneity or $\sigma_{\logit(p_{pos})}$ following [26]. We calculated the
175 within-study variance, v , [27] as:

$$176 \quad v = \frac{(k - 1) \sum_1^i \omega_i}{(\sum_1^i \omega_i)^2 - \sum_1^i \omega_i^2}$$

177 where k was the number of studies or observations included in the meta-analysis, and $\omega_i =$
178 $1/v_i$ where v_i was the variance of the proportion positive by culture, PCR or RDT within each

179 study/observation. When multiple tests were used in a study, we used the maximum variance
180 estimate across the tests.

181

182 *Data availability*

183 All input data and analytical code are available at
184 https://github.com/HopkinsIDD/cholera_positivity. This study was approved by the Johns
185 Hopkins University Institutional Review Board and Temple University Institutional Review Board.

186

187 **Results**

188 Study characteristics

189 We identified 125 studies that met our inclusion criteria (Figure 1). Of these, 113 studies
190 contained non-overlapping samples and were included in our analysis dataset (Figure 1). The
191 non-overlapping data came from 28 countries and were reported at different geographic levels,
192 including the country level (n=38 study locations) and first (n=57), second (n=76), and third
193 administrative levels (n=34) (Figure S1). Data were from 1992 through 2019 with most data
194 coming from 2015—2019 (n=67 study-periods), followed by 2010—2014 (n=45), 2005—2009
195 (n=32), and 1997—2004 (n=29) (Figure S2). Most studies were conducted in South Asia and
196 West, Central, and East Africa, with additional studies from Haiti, Yemen, Iraq, Iran, Laos,
197 Vietnam, Papua New Guinea, and the Philippines (Figure S1).

198 Most of the data were from surveillance studies (77%), followed by diagnostic test
199 accuracy studies (15%) and vaccine effectiveness studies (7%) (Table 1). Thirty percent used
200 high-quality sampling methods (i.e., tested all suspected cases, a random sample, or
201 systematically selected every nth suspected case), while the remaining 70% used convenience
202 sampling or did not report the sampling approach (Table 1). Even though most studies did not
203 include *V. cholerae* positivity disaggregated by individual-level characteristics, 28% reported the

204 proportion of suspected cases under age five, 8% reported the proportion severely dehydrated,
205 8% reported the proportion on antibiotics, and one study reported all three (Table S2).

206

207 *V. cholerae* positivity in unadjusted data

208 We found that reported *V. cholerae* positivity varied by study methodology. Median
209 positivity was 36% by culture, 37% by PCR, and 56% by RDT, with substantial overlap between
210 distributions (Figure 2A). Positivity was higher across studies that used low quality or
211 convenience sampling methods (40%) compared to those that used high quality or
212 representative sampling (33%) (Figure 2B). Positivity increased with higher minimum ages in
213 suspected cholera case definitions (Figure 2C), and we found a modest negative correlation
214 between positivity and proportion of suspected cases under age five (Spearman $r = -0.51$; 95%
215 Confidence Interval: -0.72 to -0.21) (Figure S3A).

216 Unadjusted positivity was higher when surveillance was initiated in response to an
217 outbreak (median of 45%) compared to situations where surveillance was routine or post-
218 vaccination (median of 35%) (Figure 2D). We found limited evidence for differences by mean
219 annual suspected case incidence rate in countries where these estimates were available (Figure
220 S3B), though this result should be interpreted with caution because incidence rates are for
221 2010-2016 [2] which is not the same timeframe as each study.

222 We found a modest positive correlation between positivity and the proportion of
223 suspected cases severely dehydrated (Spearman $r = 0.52$; 95% CI: 0.01 to 0.90) (Figure S3C).
224 While not statistically significant, we found a weak negative correlation between positivity and
225 the proportion of suspected cases that had received antibiotics prior to testing (Spearman $r = -$
226 0.46; 95% CI: -0.84 to 0.13) (Figure S3D).

227

228 Adjusted underlying *V. cholerae* positivity

229 Since different imperfect diagnostic tests were used to confirm *V. cholerae* O1/O139, we
230 adjusted positivity estimates from each study to account for test performance. To estimate the
231 average performance of each type of diagnostic test, we pooled estimates of sensitivity and
232 specificity across four studies that reported detailed results for all three tests (see Methods).
233 This included data from Bangladesh [24], South Sudan [16], Kenya [22], and Zambia [23]. We
234 estimated an average sensitivity of 82.0% (95% Credible Interval: 37.5-98.7) and specificity of
235 94.3% (95% CI: 81.5-99.6) for culture, an average sensitivity of 85.1% (95% CI: 53.6%-98.9%)
236 and specificity of 94.2 (95% CI: 81.8-99.7) for PCR, and an average sensitivity of 90.4% (95%
237 CI: 55.2-99.5) and specificity of 88.9% (95% CI: 54.9-99.4) for RDT (Figure 3A, Table S3).

238 After adjusting for diagnostic test performance (see Methods), we estimated that 49%
239 (95% Credible Interval: 43%-54%) of suspected cases tested were true *V. cholerae* O1/O139
240 infections (Figure 3, Figure S4, Table S4). This estimate remained constant after additionally
241 adjusting for differences in study methods (i.e., sampling quality and age minimum in case
242 definition) and outbreak setting (Figure 3, Table S4), as well as in sensitivity analysis where we
243 increased variance on all prior distributions (see Methods) (Table S4).

244 We found substantial heterogeneity between studies ($I^2 = 99.99$; $\tau^2 = 0.79$) (Figure 4).
245 When we used the model to predict positivity for a hypothetical new study site (i.e., posterior
246 predictive distribution), median positivity remained similar but, as expected, uncertainty was
247 much higher (48%, 95% CI: 7.1%-92%) (Figure S4).

248

249 Factors associated with variation in *V. cholerae* positivity

250 We then examined factors that could explain variation in *V. cholerae* positivity. After
251 adjusting for test performance, sampling quality, and outbreak setting, we found that each 1-
252 year increase in the minimum age in the case definition was associated with 1.16 (95% CI: 1.01-
253 1.34) times higher odds of a suspected cholera case having a true infection (Table S5). The
254 highest minimum age set was 10, which was used in a single study (Figure 2C). We observed a

255 similar trend when we analyzed minimum age in case definition as a categorical rather than
256 continuous variable, though the increases in odds of positivity by age in definition were not
257 significant and had high uncertainty (Table S6).

258 We estimated that the odds of a suspected cholera case having a true *Vibrio cholerae*
259 O1/O139 infection were 1.64 (95% CI: 1.06-2.52) times higher when surveillance was initiated in
260 response to a cholera outbreak compared to non-outbreak surveillance, after adjusting for test
261 performance, sampling quality, and case definition (Table S5). These estimates remained
262 consistent across sensitivity analyses (Table S7).

263

264 Discussion

265 Here we estimated that on average half of medically attended suspected cholera cases
266 represent true *V. cholerae* O1/O139 infections. We found that *V. cholerae* positivity was higher
267 when higher ages were set in case definitions and when surveillance was initiated in response
268 to an outbreak. Additionally, we found substantial heterogeneity in *V. cholerae* positivity
269 between studies, so that simply dividing suspected cholera case counts by two to estimate the
270 true number of cases will not be appropriate in all settings.

271 These findings have several implications for cholera surveillance policy. The GTFCC
272 defines suspected cholera in areas where an outbreak has not yet been reported as acute
273 watery diarrhea and severe dehydration or death in individuals two years and older [17]. Our
274 finding that a minimum age of two or five increases specificity for identifying a true *V. cholerae*
275 infection in suspected cases supports this case definition. Laboratory confirmation tests are
276 recommended during the early stages of an outbreak and periodically once an outbreak has
277 been declared, ideally a minimum of five suspected case samples tested per week per health
278 facility [17]. Our finding that high quality sampling also increases specificity for *V. cholerae*
279 suggests that systematically selecting cases to test is important for accurately evaluating
280 endemic cholera. We also found high variability across settings, suggesting that when

281 systematic testing is not possible, assumptions about positivity based on data at the location
282 and time where cholera control decisions are made may be key to reducing bias in estimates of
283 true cholera incidence.

284 A remaining question is why only about half of medically attended suspected cholera
285 cases represent true infections. It is possible that we overestimated test sensitivity and have not
286 fully accounted for false negatives; unfortunately, this is difficult to evaluate without a gold
287 standard diagnostic test. A portion of the remaining suspected cases could also be infections
288 with other enteric pathogens, especially those with similar transmission modes as cholera that
289 may have outbreaks or high levels of endemic transmission concurrently. For example, in Uvira,
290 Democratic Republic of the Congo, 36% of suspected cholera cases were positive for
291 Enterotoxigenic *Escherichia coli* and 28% for *Cryptosporidium* [10]. In rural Bangladesh, the
292 majority of acute watery diarrhea in children under 18 months was attributable to rotavirus, while
293 older children were more often infected with *V. cholerae* [12]. In Haiti, 64% of acute watery
294 diarrhea cases tested positive for *V. cholerae* O1, 4% for rotavirus, and <1% for Shigella and
295 Salmonella, though rotavirus positivity was higher among children under five [11]. Thus, the
296 relative contribution of non-cholera watery diarrhea varies with age distribution and other
297 location-specific drivers of enteric infections.

298 One of the limitations of this study was that we could not account for all potential drivers
299 of *V. cholerae* positivity, which contributed to the large heterogeneity we found between studies.
300 In addition, *V. cholerae* positivity may be highest in the early stages of an outbreak [7,9,28], but
301 we could not account for this given the temporal resolution of our dataset. However, a strength
302 of our approach is that we pooled estimates from studies across diverse geographies, time
303 periods, and epidemiological contexts. A further potential limitation is that, without a gold
304 standard diagnostic test, sensitivity and specificity estimates may be biased if the tests are less
305 sensitive and/or specific for shared reasons. The hierarchical conditional dependence model we
306 used accounted for this pairwise dependence and increased uncertainty around our estimates

307 accordingly. This approach also allowed us pool test performance estimates across studies from
308 four countries. Thus, to our knowledge, we adjusted our estimates for test sensitivity and
309 specificity using the best generic estimates available. Overall, we have high confidence in our
310 average estimates of *V. cholerae* positivity, despite the difficulty of accurately estimating
311 positivity in a new location/time without confirmation tests.

312 These estimates of *V. cholerae* positivity address one part of the challenge in
313 establishing the true burden of cholera: cases that are overcounted due to non-specific
314 suspected case definitions. A crucial next step will be to estimate missed cases due to care
315 seeking and poor clinical surveillance. This could be done in part through systematically
316 synthesizing data from studies of care seeking behavior for diarrheal symptoms (for example
317 [29,30]), including where potential cholera cases seek care (e.g., at pharmacies, traditional
318 healers, or hospitals). This could additionally be done through population representative surveys
319 and active case finding, similar to studies conducted in Haiti [31] and Tanzania [32],
320 respectively, which demonstrated higher mortality rates associated with cholera than had been
321 reported through passive surveillance.

322 Ultimately, a better understanding of *V. cholerae* positivity will help us move towards
323 estimates of true cholera incidence and mortality. Given the large amount of heterogeneity
324 between studies, it will be important to do this in a way that accounts for variation in *V. cholerae*
325 positivity between sites. Moreover, the proportion of suspected cholera cases missed because
326 of milder symptoms or barriers to healthcare seeking needs to be estimated and accounted for.
327 Such estimates will provide crucial information to guide the allocation of limited resources such
328 as vaccines in a way that most effectively supports cholera prevention and control.

329

330 **Funding**

331 This work was supported by the Bill and Melinda Gates Foundation [grant number
332 OPP1171700] and the National Institute of Allergy and Infectious Disease [grant number

333 AI135115-01A1].

334

335 **Conflicts of Interest**

336 The authors declare no conflicts of interest.

337

338 **References**

- 339 1. Ali M, Nelson AR, Lopez AL, Sack DA. Updated global burden of cholera in endemic
340 countries. *PLOS Neglected Tropical Diseases* **2015**; 9:e0003832.
- 341 2. Lessler J, Moore SM, Luquero FJ, et al. Mapping the burden of cholera in sub-Saharan
342 Africa and implications for control: an analysis of data across geographical scales. *The*
343 *Lancet* **2018**; 391:1908–1915.
- 344 3. Azman AS, Lauer SA, Bhuiyan TR, et al. *Vibrio cholerae* O1 transmission in Bangladesh:
345 insights from a nationally representative serosurvey. *The Lancet Microbe* **2020**; 1:e336–
346 e343.
- 347 4. Mosley WH, Benenson AS, Barui R. A serological survey for cholera antibodies in rural
348 East Pakistan. *Bull World Health Organ* **1968**; 38:327–334.
- 349 5. Weil AA, Begum Y, Chowdhury F, et al. Bacterial shedding in household contacts of
350 cholera patients in Dhaka, Bangladesh. *Am J Trop Med Hyg* **2014**; 91:738–742.
- 351 6. Khan AI, Rashid MM, Islam MT, et al. Epidemiology of cholera in Bangladesh: findings
352 from nationwide hospital-based surveillance, 2014–2018. *Clinical Infectious Diseases*
353 **2020**; 71:1635–1642.
- 354 7. Camacho A, Bouhenia M, Alyusfi R, et al. Cholera epidemic in Yemen, 2016–18: an
355 analysis of surveillance data. *The Lancet Global Health* **2018**; 6:e680–e690.

- 356 8. Sauvageot D, Njanpop-Lafourcade B-M, Akilimali L, et al. Cholera incidence and mortality
357 in sub-Saharan African sites during multi-country surveillance. *PLOS Neglected Tropical*
358 *Diseases* **2016**; 10:e0004679.
- 359 9. Jameel SK, Shafek MA, Abdulmohsen AM, Mohamed NS, Najji SR, Mohammed TT. The
360 isolation of *Vibrio cholera* and other enteric bacteria with molecular characterization of
361 *Vibrio cholera* during the outbreak of Baghdad/Iraq in 2015. *Advances in Microbiology*
362 **2016**; 6:699–715.
- 363 10. Williams C, Cumming O, Grignard L, et al. Prevalence and diversity of enteric pathogens
364 among cholera treatment centre patients with acute diarrhea in Uvira, Democratic Republic
365 of Congo. *BMC Infectious Diseases* **2020**; 20:741.
- 366 11. Steenland MW, Joseph GA, Lucien MAB, et al. Laboratory-confirmed cholera and rotavirus
367 among patients with acute diarrhea in four hospitals in Haiti, 2012–2013. *Am J Trop Med*
368 *Hyg* **2013**; 89:641–646.
- 369 12. Siddique AK, Ahmed S, Iqbal A, et al. Epidemiology of rotavirus and cholera in children
370 aged less than five years in rural Bangladesh. *J Health Popul Nutr* **2011**; 29:1–8.
- 371 13. Nadri J, Sauvageot D, Njanpop-Lafourcade B-M, et al. Sensitivity, specificity, and public-
372 health utility of clinical case definitions based on the signs and symptoms of cholera in
373 Africa. *The American Journal of Tropical Medicine and Hygiene* **2018**; 98:1021–1030.
- 374 14. Lucien MAB, Schaad N, Steenland MW, et al. Identifying the most sensitive and specific
375 sign and symptom combinations for cholera: results from an analysis of laboratory-based
376 surveillance data from Haiti, 2012–2013. *The American Journal of Tropical Medicine and*
377 *Hygiene* **2015**; 92:758–764.

- 378 15. Muzembo BA, Kitahara K, Debnath A, Okamoto K, Miyoshi S-I. Accuracy of cholera rapid
379 diagnostic tests: a systematic review and meta-analysis. *Clinical Microbiology and Infection*
380 **2021**; 0. Available at: <https://www.clinicalmicrobiologyandinfection.com/article/S1198->
381 [743X\(21\)00489-4/fulltext](https://www.clinicalmicrobiologyandinfection.com/article/S1198-743X(21)00489-4/fulltext). Accessed 2 November 2021.
- 382 16. Ontweka LN, Deng LO, Rauzier J, et al. Cholera rapid test with enrichment step has
383 diagnostic performance equivalent to culture. *PLOS ONE* **2016**; 11:e0168257.
- 384 17. Global Task Force on Cholera Control (GTFCC) Surveillance Working Group. Interim
385 guidance document on cholera surveillance. 2017; Available at: [https://www.gtfcc.org/wp-](https://www.gtfcc.org/wp-content/uploads/2019/10/gtfcc-interim-guidance-document-on-cholera-surveillance.pdf)
386 [content/uploads/2019/10/gtfcc-interim-guidance-document-on-cholera-surveillance.pdf](https://www.gtfcc.org/wp-content/uploads/2019/10/gtfcc-interim-guidance-document-on-cholera-surveillance.pdf).
387 Accessed 30 September 2022.
- 388 18. Weill F-X, Domman D, Njamkepo E, et al. Genomic history of the seventh pandemic of
389 cholera in Africa. *Science* **2017**; 358:785–789.
- 390 19. Wang C, Lin X, Nelson KP. Bayesian hierarchical latent class models for estimating
391 diagnostic accuracy. *Stat Methods Med Res* **2020**; 29:1112–1128.
- 392 20. Plummer M, Stukalov A, Denwood M. rjags: Bayesian graphical models using MCMC.
393 2022; Available at: <https://CRAN.R-project.org/package=rjags>. Accessed 17 August 2022.
- 394 21. R Core Team. R: A language and environment for statistical computing. R Foundation for
395 Statistical Computing, Vienna, Austria. [http://wwwR-project.org/](http://www.R-project.org/) **2016**; Available at:
396 <https://cir.nii.ac.jp/crid/1574231874043578752>. Accessed 17 August 2022.
- 397 22. Debes AK, Murt KN, Waswa E, et al. Laboratory and field evaluation of the Crystal VC-O1
398 cholera rapid diagnostic test. *The American Journal of Tropical Medicine and Hygiene*
399 **2021**; 104:2017–2023.

- 400 23. Mwaba J, Ferreras E, Chizema-Kawesa E, et al. Evaluation of the SD bioline cholera rapid
401 diagnostic test during the 2016 cholera outbreak in Lusaka, Zambia. *Tropical Medicine &*
402 *International Health* **2018**; 23:834–840.
- 403 24. Sayeed MA, Islam K, Hossain M, et al. Development of a new dipstick (Cholkit) for rapid
404 detection of *Vibrio cholerae* O1 in acute watery diarrheal stools. *PLOS Neglected Tropical*
405 *Diseases* **2018**; 12:e0006286.
- 406 25. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* **2002**;
407 21:1539–1558.
- 408 26. Gronau QF, Heck DW, Berkhout SW, Haaf JM, Wagenmakers E-J. A primer on Bayesian
409 model-averaged meta-analysis. *Advances in Methods and Practices in Psychological*
410 *Science* **2021**; 4:25152459211031256.
- 411 27. Viechtbauer W. I2 for multilevel and multivariate models. 2022. Available at:
412 https://www.metafor-project.org/doku.php/tips:i2_multilevel_multivariate. Accessed 26
413 August 2022.
- 414 28. Zgheir SM, Mustafa NM, Ali AA, Al-Diwan J. Cholera Outbreak in Iraq, 2017. *Ind Jour of*
415 *Publ Health Rese & Develop* **2019**; 10:686.
- 416 29. Chowdhury F, Khan IA, Patel S, et al. Diarrheal illness and healthcare seeking behavior
417 among a population at high risk for diarrhea in Dhaka, Bangladesh. *PLOS ONE* **2015**;
418 10:e0130105.
- 419 30. Fissehaye T, Damte A, Fantahun A, Gebrekirstos K. Health care seeking behaviour of
420 mothers towards diarrheal disease of children less than 5 years in Mekelle city, North
421 Ethiopia. *BMC Res Notes* **2018**; 11:749.

- 422 31. Luquero FJ, Rondy M, Boncy J, et al. Mortality rates during cholera epidemic, Haiti, 2010–
423 2011. *Emerging Infectious Diseases* **2016**; 22. Available at:
424 https://wwwnc.cdc.gov/eid/article/22/3/14-1970_article. Accessed 9 September 2022.
- 425 32. McCrickard LS, Massay AE, Narra R, et al. Cholera mortality during urban epidemic, Dar
426 es Salaam, Tanzania, August 16, 2015–January 16, 2016. *Emerging Infectious Diseases*
427 **2017**; 23. Available at: https://wwwnc.cdc.gov/eid/article/23/13/17-0529_article. Accessed
428 9 September 2022.
- 429
- 430

431 **Tables**

432 **Table 1. Study characteristics**

433 Number of study-observations included in the dataset with each study characteristic. An individual study
 434 has >1 study-observations if it reports *V. cholerae* positivity for >1 sampling method or country.

Category	Characteristic	Number of studies	Percent of studies
Study design	Surveillance	96	76.8
	Diagnostic test accuracy	19	15.2
	Vaccine effectiveness	9	7.2
	Other	1	0.8
Sampling method quality	High	37	29.6
	Low	88	70.4
Percent of suspected cases tested	0-4	15	12.0
	5-49	31	24.8
	50-95	22	17.6
	≥95	57	45.6
Number of tests used (of culture, PCR, and/or RDT)	1	98	78.4
	2	20	16.0
	≥3	7	5.6
Number of suspected cases tested	1-9 [†]	1	0.8
	10-99	37	29.6
	100-999	62	49.6
	≥1000	25	20.0

435 [†]One multi-country surveillance study overall tested ≥10 suspected cholera cases for
 436 *V. cholerae* O1/O139 but reported fewer than 10 tested in one country.

437 **Figure legends**

438 **Figure 1. PRISMA flow diagram**

439 Diagram illustrating literature selection process, including databases searched, literature
440 screened, and full texts reviewed for eligibility. Reasons for exclusion are indicated along with
441 the number of studies that fell within each category.

442

443 **Figure 2. *Vibrio cholerae* positivity by study methodology and outbreak context**

444 Proportion of suspected cholera cases that were confirmed positive by **A)** diagnostic test type,
445 **B)** quality of sampling methods, where "high" includes all suspected cases or a random or
446 stratified sample and "low" includes convenience or unreported sampling methods, **C)** age
447 minimum in suspected case definition, where "0" indicates that no minimum age was set, and **D)**
448 whether surveillance was initiated in response to an outbreak or whether it was routine
449 surveillance or non-outbreak. Each point is a study-observation; individual studies may have
450 multiple points if they reported positivity by multiple tests, reported results for multiple countries
451 or outbreak contexts, and/or multiple sampling methods. Circles represent study-observations
452 with high quality sampling and triangles indicate low quality sampling.

453

454 **Figure 3. Estimated underlying *V. cholerae* positivity**

455 **A)** Posterior distributions of pooled percent sensitivity and specificity of culture (top), PCR
456 (middle), and RDT (bottom) for detecting *V. cholerae* O1/O139 infections in suspected cholera
457 cases. Dashed lines represent median values of each distribution. **B)** "Unadjusted" is mean
458 percent positive (95% confidence interval) from random effects meta-analysis implemented
459 using R package meta without any adjustments. "Adjusted" is the underlying true percent
460 positive (95% credible interval) estimated using the Bayesian hierarchical model, adjusted for
461 sensitivity/specificity of the tests, minimum age in case definition, sampling quality, and whether
462 surveillance was initiated in response to an outbreak.

463

464 **Figure 4. Forest plot of study estimates and underlying positivity**

465 Black points indicate median study-level underlying positivity and 95% Credible Interval (CI).

466 Teal, orange, and purple points indicate the proportion positive reported by study for culture,

467 PCR, and RDT, respectively, and corresponding error bars indicate 95% confidence interval for

468 a binomial probability. Studies are labeled by country ISO3 code, whether they used high quality

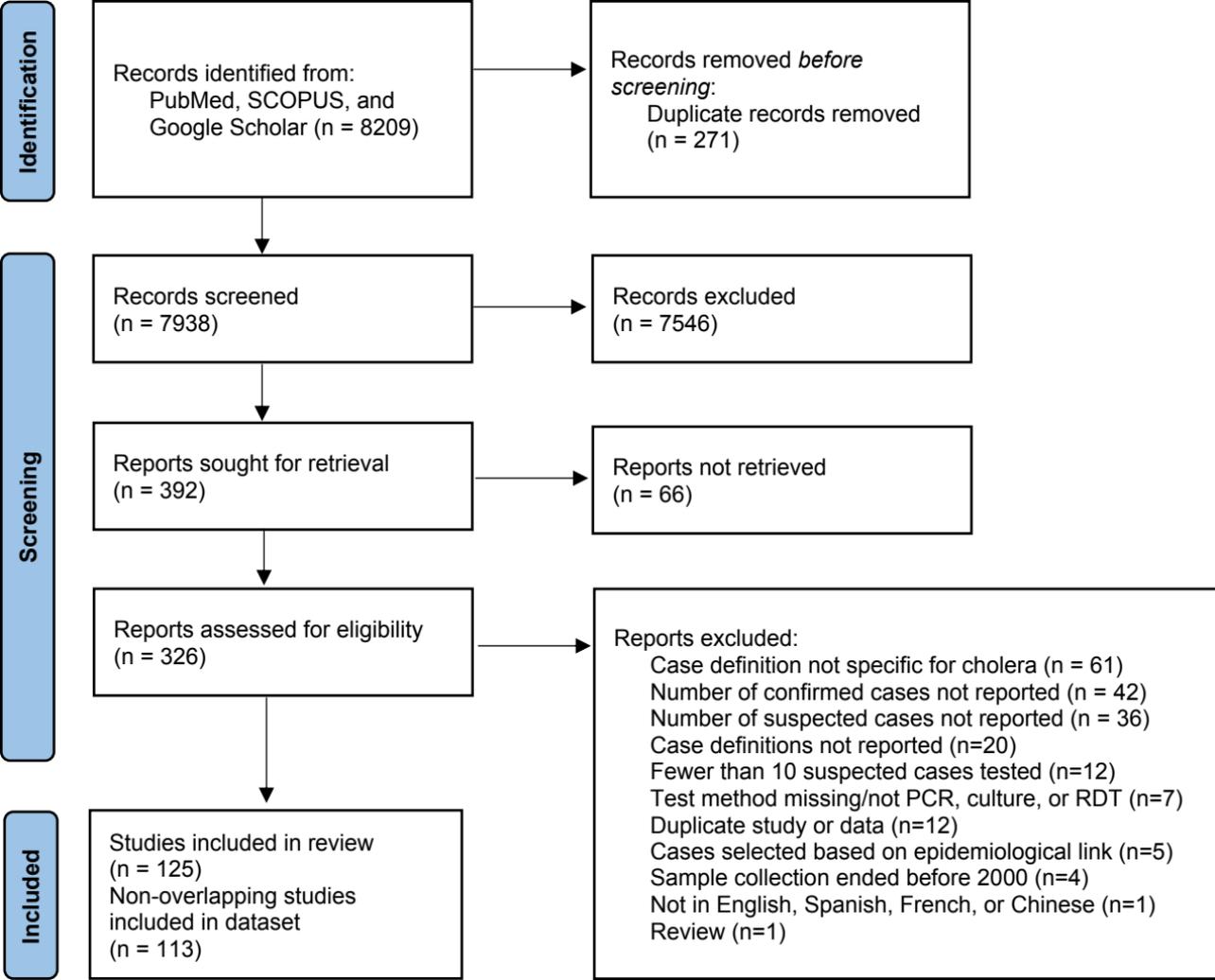
469 sampling methods, and whether a minimum age was set in the suspected cholera case

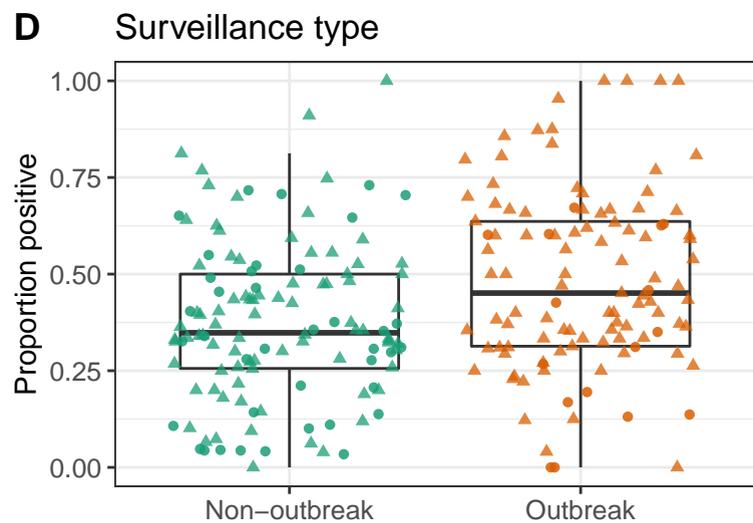
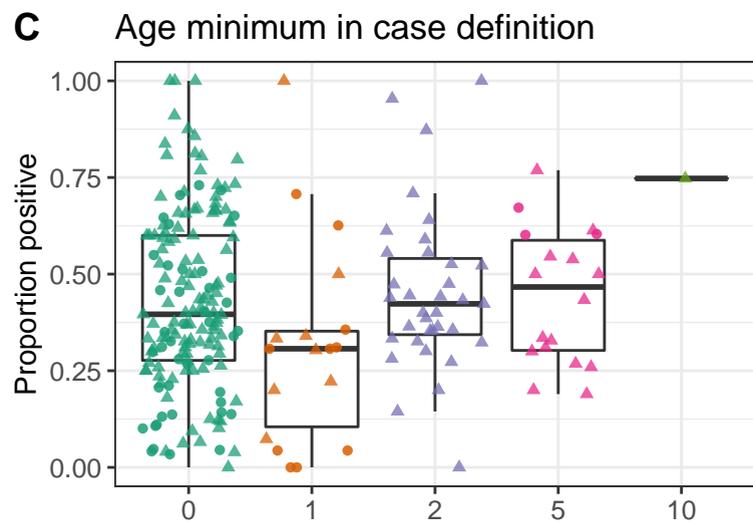
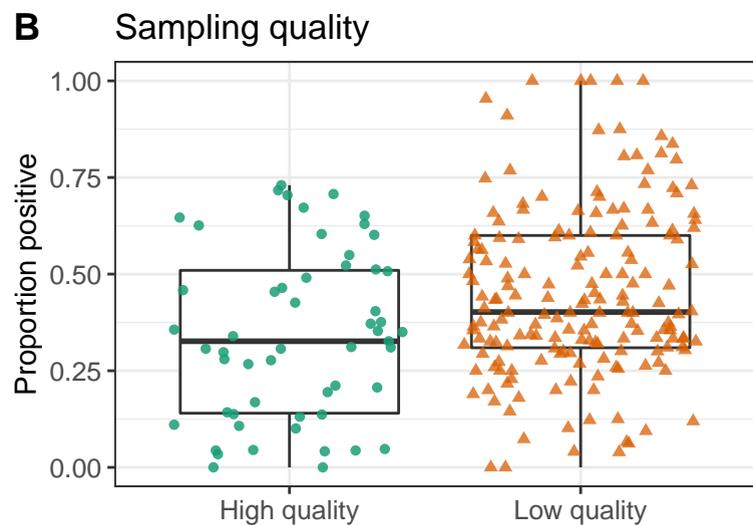
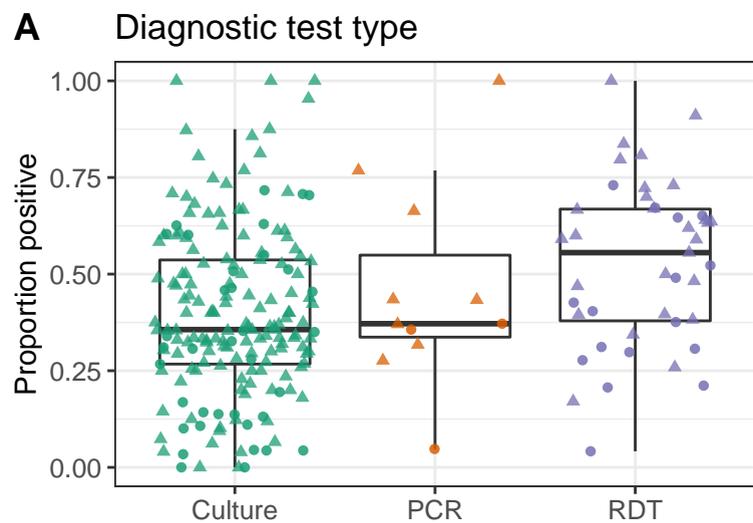
470 definition. Green distribution indicates median and 95% CI of pooled positivity. Pink distribution

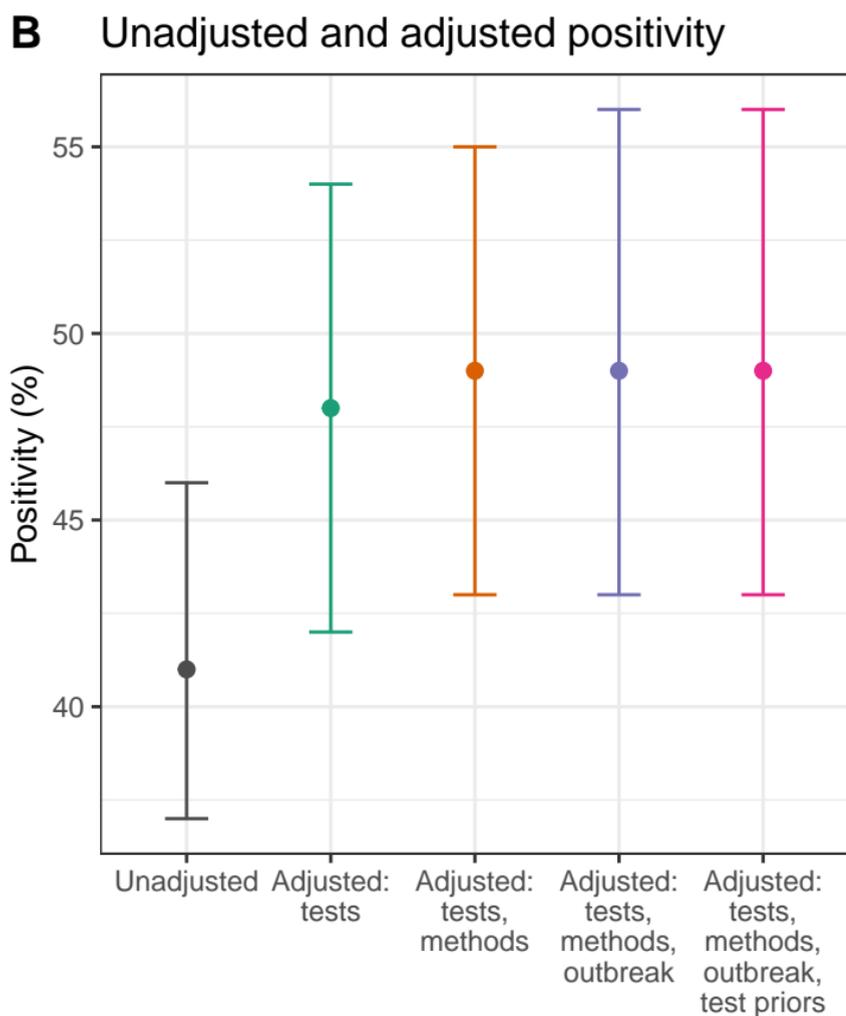
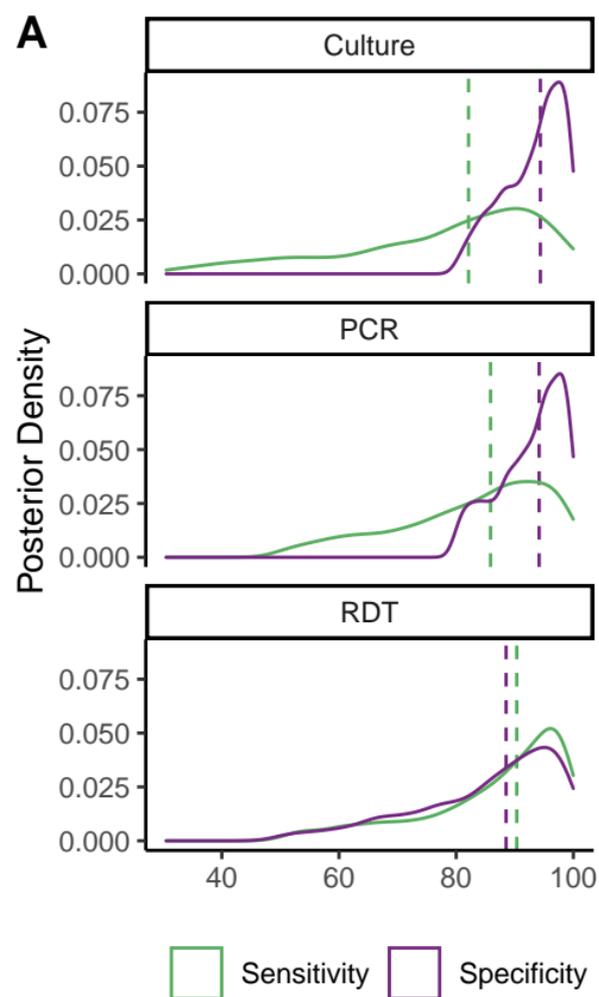
471 indicates median and 95% CI of posterior predictive distribution of positivity. Studies are split

472 into outbreak and non-outbreak for ease of interpretation, but the green and pink distributions

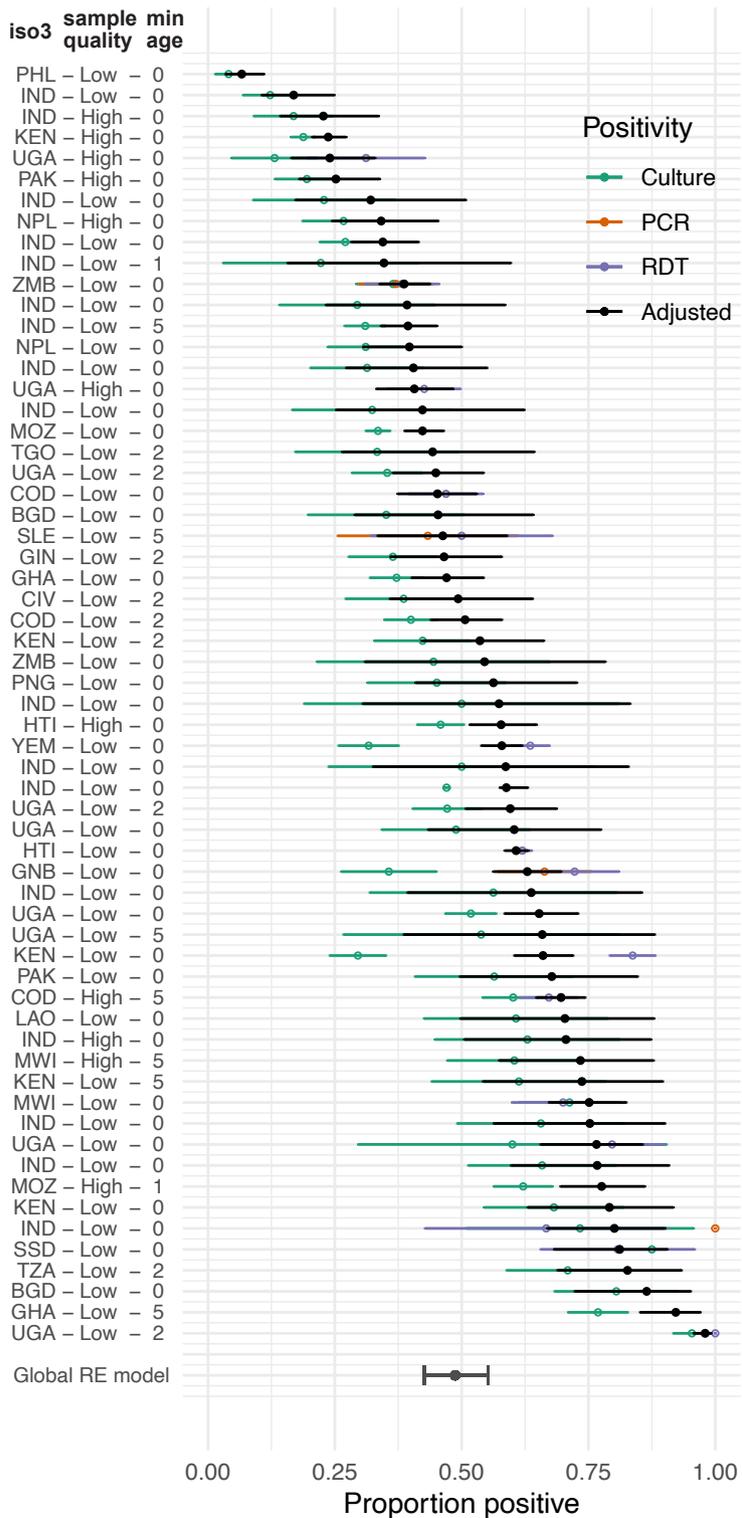
473 represent the global estimates across all studies.







Outbreak surveillance



Non-outbreak surveillance

