

1 **A general framework to support cost-efficient survey design**
2 **choices for the control of soil-transmitted helminths when**
3 **deploying Kato-Katz thick smear**

4
5 **Running title:** Survey designs for STH control programs

6
7 Adama Kazienga^{1,2*}, Bruno Levecke¹, Gemechu Tadesse Leta³, Sake J. de Vlas², Luc
8 E. Coffeng²

9
10 ¹Department of Translational Physiology, Infectiology and Public Health, Ghent
11 University, Merelbeke, Belgium

12 ²Department of Public Health, Erasmus MC, University Medical Center Rotterdam,
13 Rotterdam, The Netherlands

14 ³Bacterial, Parasitic and Zoonotic Diseases Research Directorate, Ethiopian Public
15 Health Institute, Addis Ababa, Ethiopia

16
17 *Corresponding author: kazienga_adama@yahoo.fr (KA).

18

19

20

21

22

23

24

25 **Abstract**

26 **Background**

27 To monitor and evaluate soil-transmitted helminth (STH) control programs, the World
28 Health Organization (WHO) recommends screening stools from 250 children across 5
29 schools, deploying Kato-Katz thick smear (KK). However, it remains unclear whether
30 these recommendations are sufficient to make adequate decisions about stopping
31 preventive chemotherapy (PC) (prevalence of infection <2%) or declaring elimination
32 of STH as a public health problem (prevalence of moderate-to-heavy intensity (MHI)
33 infections <2%).

34

35 **Methodology**

36 We developed a simulation framework to determine the effectiveness and cost of
37 survey designs for decision-making in STH control programs, capturing the
38 operational resources to perform surveys, the variation in egg counts across STH
39 species, across schools, between and within individuals, and between repeated
40 smears. Using this framework and a lot quality assurance sampling approach, we
41 determined the most cost-efficient survey designs (number of schools, subjects, stool
42 samples per subject, and smears per stool sample) for decision-making.

43

44 **Principal findings**

45 For all species, employing duplicate KK (sampling 4 to 6 schools and 64 to 70 subjects
46 per school) was the most cost-efficient survey design to assess whether prevalence
47 of any infection intensity was above or under 2%. For prevalence of MHI infections,
48 single KK was the most cost-efficient (sampling 11 to 25 schools and 52 to 84 children
49 per school).

50 **Conclusions/Significance**

51 KK is valuable for monitoring and evaluation of STH control programs, though we
52 recommend to deploy a duplicate KK on a single stool sample to stop PC, and a single
53 KK to declare the elimination of STH as a public health problem.

54

55 **Keywords:** soil-transmitted helminths, Kato-Katz thick smear, lot quality assurance
56 sampling, survey design, total operational cost

57 **Author summary**

58 Worldwide large-scale deworming programs are implemented to reduce the morbidity
59 attributable to intestinal worms in school children. To monitor and evaluate the
60 progress towards the programs goals, World Health Organization (WHO) has
61 developed both a survey design and a corresponding decision tree based on the
62 prevalence and intensity of infections. However, these programs operate in resource-
63 constrained countries, and hence it is crucial to minimize the operational costs to
64 survey worm infections while ensuring the correctness of the program decision. To
65 further support WHO in more evidence-based recommendations for cost-efficient
66 decision-making, we developed a general framework that captures both the
67 operational resources to perform surveys and the variation in test results when
68 deploying the current diagnostic standard. Subsequently, we determined the most
69 cost-efficient survey design to decide to stop the deworming programs and to verify
70 whether the morbidity attributable to intestinal worms has been eliminated as a public
71 health problem. Generally, we found that the current WHO-recommended survey
72 design may not allow for optimal decision making. Based on our results, we proposed
73 alternative survey designs for each of the worm species and program targets.

74

75

76

77

78

79

80 Introduction

81 Recently, the World Health Organization (WHO) released its 2030 roadmap for the
82 control of soil-transmitted helminths (STHs) [1,2]. In this new roadmap, the two crucial
83 targets are (i) to eliminate STHs as a public health problem (EPHP), defined as
84 prevalence of moderate-to-heavy intensity (MHI) infections less than 2%, and (ii) to
85 reduce the number of tablets needed in preventive chemotherapy (PC). Regarding this
86 second target, WHO recommends to stop PC if the prevalence of infection (any
87 intensity) is less than 2%.

88

89 To assess whether prevalence of any intensity or MHI infection is under 2%, the WHO
90 recommends screening stool samples from 250 children across 5 schools (50 children
91 per school), using single Kato-Katz thick smears (KK) [2]. However, little attention has
92 been paid to whether this survey design is sufficient for reliable decision-making.
93 There will always be a risk of making a wrong decision: stopping PC or declaring EPHP
94 too early (referred to as “undertreatment” from here on) or too late (“overtreatment”).
95 Amongst other things, these risks are driven by survey sample size and diagnostic
96 accuracy, which means that a trade-off has to be made between operational costs and
97 feasibility of particular survey designs, and the maximum acceptable risks of under-
98 and overtreatment.

99

100 Here, we assess the trade-off of survey design and associated operational costs
101 versus the risk of under- and overtreatment for STHs when using KK as a diagnostic
102 method because it was found more cost-effective than alternative diagnostic
103 techniques based on egg counting [3]. We do this using a previously developed
104 simulation framework for lot quality assurance sampling (LQAS) that only captured the

105 prevalence of infection and not the variation in intensity of infection [4]. For the purpose
106 of this study, we expanded this framework to also capture variation in intensity of
107 infection across STH species (*Ascaris lumbricoides*, *Trichuris trichiura*, hookworms),
108 across schools, between and within individuals, and between repeated thick smears.
109 These sources of variation are important as the sensitivity of KK for detecting infection
110 in an individual depends on the intensity of infection [5–9]. Furthermore, using this
111 expanded framework, we assess to what extent the accuracy of survey results can be
112 improved (and at what cost) by increasing the number of thick smears per stool sample
113 or the number of stool samples per person [5,10]. Finally, for different levels of
114 acceptable risk of under- and overtreatment, we determine the most cost-efficient
115 survey design and the associated decision cut-off (maximum number of egg-positive
116 individuals) for decision-making.

117

118 **Methods**

119 **Overview**

120 Based on a previously developed LQAS framework for monitoring and evaluation
121 (M&E) of neglected tropical diseases (NTDs) [4], we developed a generic simulation
122 framework for STH surveys and the associated operational costs at the level of a PC
123 implementation unit. Using this framework, we assessed the cost-efficiency of different
124 survey designs (number of schools, number of children per school, number of stool
125 samples per child, and number of KK thick smears per stool sample) for each STH
126 species and decision type (stop PC or declare EPHP) via the following steps: (1)
127 simulate egg counts in thick smears of sampled children and determine the number of
128 children that test positive; (2) for all possible decision cut-offs in terms of maximum

129 number of positive children, calculate the probability of overtreatment and
130 undertreatment across repeated Monte Carlo simulations and determine the decision
131 cut-off for adequate decision-making (acceptable risk of under- and overtreatment);
132 (3) estimate the total survey cost. To determine the most cost-efficient survey design
133 and associated cut-off, we selected all designs that resulted in adequate decision-
134 making and compared them in terms of cost and feasibility to implement.

135

136 **Simulation of egg counts for different survey designs**

137 We adopted the simulation framework for egg counts developed by Coffeng et al. [11]
138 and expanded it with geographical variation in infection levels. The expanded
139 framework simulates egg count data from a compound lognormal-gamma-gamma-
140 gamma-Poisson distribution, capturing the following sources of variability in egg
141 counts:

- 142 1. Variability in mean egg per gram stool (EPG) between schools (assumed to be
143 lognormal distributed (see **Supplement S1**);
- 144 2. Inter-individual variability in EPG due to variation in infection levels between
145 individuals (considered to be gamma-distributed), where the level of
146 aggregation (shape of the gamma distribution) is a linear function of the school-
147 level mean EPG (see **Supplement S1**);
- 148 3. Day-to-day variability in mean EPG within an individual due to heterogeneous
149 egg shedding over time (assumed to be gamma-distributed);
- 150 4. Variability in mean EPG between repeated slides based on the same sample
151 due to the aggregated distribution of eggs in feces (assumed to follow a gamma
152 distribution);

153 5. Variability in count observation (expressed in raw egg count) due to random
154 diagnostic variation (assumed to be Poisson distributed).

155
156 We presented the mathematical backbone of the simulation framework in **Supplement**
157 **S2** and quantified these aforementioned sources of variation based on published
158 datasets [5,12]. **Table 1** provides the full overview of the estimated parameter values
159 for each STH species.

160
161 **Table 1. Parametrization of the simulation framework for various sources of variability in Kato-**
162 **Katz thick smear egg counts.** We parametrized the lognormal distribution using the mean and
163 standard deviation on the logarithmic scale, and for the gamma distribution, the shape parameter k and
164 rate ($\frac{\mu}{k}$), where μ is the distribution's mean. To quantify k , the coefficient of variation (cv) was used as
165 standardized measure of variability [13].

166

Parameters	STH species			Data source
	<i>Ascaris</i>	Hookworm	<i>Trichuris</i>	
Variability in mean EPG across schools within the same district (σ_i)	0.69	0.80	0.52	[12]
Intercept for school-level aggregation parameter (β_0) as a linear function of the school-level mean EPG	0.0158	0.0162	0.0098	[12]
Slope for school-level aggregation parameter (β_1) as a linear function of the school-level mean EPG	0.0019	0.0222	0.0444	[12]
Day-to-day variation in EPG within an individual (shape k_d , expressed as $\frac{1}{cv^2}$)	0.5102	0.8734	1.4172	[3]

167

168

169 Using this simulation framework, we considered 4 survey designs, denoted as $KK_{a \times b}$,
170 a is the number of stool samples per person, and b is the number of repeated smears
171 per stool sample. These survey designs were $KK_{1 \times 1}$ (considered as the reference
172 [1,2]), $KK_{1 \times 2}$, $KK_{2 \times 1}$, and $KK_{2 \times 2}$. We further set the number of schools ($n_{schools}$)
173 ranging from 3 to 10 for any intensity of infection and from 10 to 25 for MHI infections,
174 while the number of children per school ($n_{children}$) ranged from 10 to 200 for both
175 targets.

176

177 **Calculating the probability of under- and overtreatment**

178 To determine the probability of making right and wrong policy decisions for different
179 survey designs, we used a 2-stage LQAS approach, a framework for NTD control
180 program decision-making which we have described in detail previously [4]. The 2-
181 stage aspect refers to the fact that we explicitly consider that survey results will be
182 sampled from multiple clusters. In 2-stage LQAS, decision-making is based on the
183 total number of positive test results (X_+ ; here, egg-positive individuals) in a sample of
184 pre-defined size from a pre-defined number of clusters in an implementation unit and
185 a decision cut-off c . The frequency of PC will be reduced (or PC is stopped altogether)
186 in case X_+ is less than the decision cut-off c . Otherwise, the frequency of PC remains
187 unchanged or may even be increased. This framework was used to assess two
188 features of the decision: (1) the operational program decision threshold (e.g., 2%
189 prevalence of infection); (2) the maximum acceptable risk of making a wrong decision
190 when the true prevalence based on a single KK is at a pre-defined lower or higher
191 value than the operational decision threshold (i.e., the “grey zone”, e.g., 1% and 3%).

192

193 Using LQAS, we assessed two operational decision thresholds for each STH species:
 194 one for a 2% prevalence of any intensity of infection (i.e., related to stopping PC) and
 195 another for a 2% prevalence of MHI infections (i.e., related to declare EPHP). Then,
 196 we considered adequate decision-making by setting the maximum allowed probability
 197 of unnecessarily continuing with the program ($E_{overtreat}$) when true prevalence is 3%
 198 (the upper limit of the grey zone, or UL) at 0.25. We further set the highest allowed
 199 probability of prematurely stopping interventions ($E_{undertreat}$) when true prevalence is
 200 1% (the lower limit of the grey zone, or LL) at 0.05. For this, we determined the mean
 201 EPG at the implementation unit level that corresponds to these limits (LL and UL) as
 202 measured by single KK (**Table 2**), conditional on the estimated parameter values for
 203 sources of variability in egg counts (**Table 1**), assuming 100% specificity of KK [14,15].

204

205 **Table 2. Parametrization of simulations for 2-stage lot quality assurance sampling in terms of**
 206 **the natural logarithm of the geometric mean infection intensity (eggs per gram feces or EPG) in**
 207 **the population as measured by single Kato-Katz.**

208

Parameters	STH species		
	<i>Ascaris</i>	Hookworm	<i>Trichuris</i>
<i>Any intensity infections</i>			
Lower limit mean EPG	log(0.449)	log(0.280)	log(0.298)
Upper limit mean EPG	log(2.691)	log(1.013)	log(0.980)
<i>Moderate-to-heavy intensity infections</i>			
Lower limit mean EPG	log(314)	log(164)	log(145)
Upper limit mean EPG	log(620)	log(287)	log(217)

209

210

211 **Calculating the operational costs and determining the most cost-efficient** 212 **survey design**

213 We estimated the total survey costs for each simulated survey design and STH
214 species. This cost was composed of (i) the cost of consumables to collect and process
215 samples, (ii) the personnel costs, and (iii) the travel costs [4]. Technical details on
216 how the total survey costs were estimated can be found in **Supplement S3**.

217

218 To determine the most cost-efficient survey design and associated decision cut-off,
219 we selected all designs that allowed for adequate decision-making. Then, we chose
220 the survey design that resulted in the lowest total survey cost (C_{tot}) and required a
221 maximum of 100 children per school (a reasonable maximum size of individuals in a
222 school).

223 **Sensitivity analysis**

224 We assessed the impact of several critical parameters on the estimated cost-efficiency
225 of different survey designs. First, we explored the impact of assuming that variation
226 between individuals within schools is governed by a fixed aggregation parameter k_k
227 (in contrast to letting it vary as a linear function of the mean EPG in a school). Second,
228 we assessed the impact of considering higher inter-individual variability in egg counts
229 (k_k) in (post-) control scenarios due to incomplete deworming coverage compared to
230 this variability in the main analysis (assumed to be as observed in pre-control settings).
231 Therefore, we set k_k to 2/3 of the value in the main analysis. Third, because diagnostic
232 performance impacts survey design [4,16], we evaluated the impact of KK specificity,
233 assuming 99.5% specificity instead of 100% in the main analysis. Finally, to assess

234 the impact of stricter choices about the maximum allowed probability of making
235 incorrect decisions, we set the risk of undertreating ($E_{undertreat}$) to 0.025 instead of
236 0.05 in the main analysis. As a second step, we also decreased the risk of overtreating
237 ($E_{overtreat}$) from 0.25 to 0.1.

238

239

240 **Results**

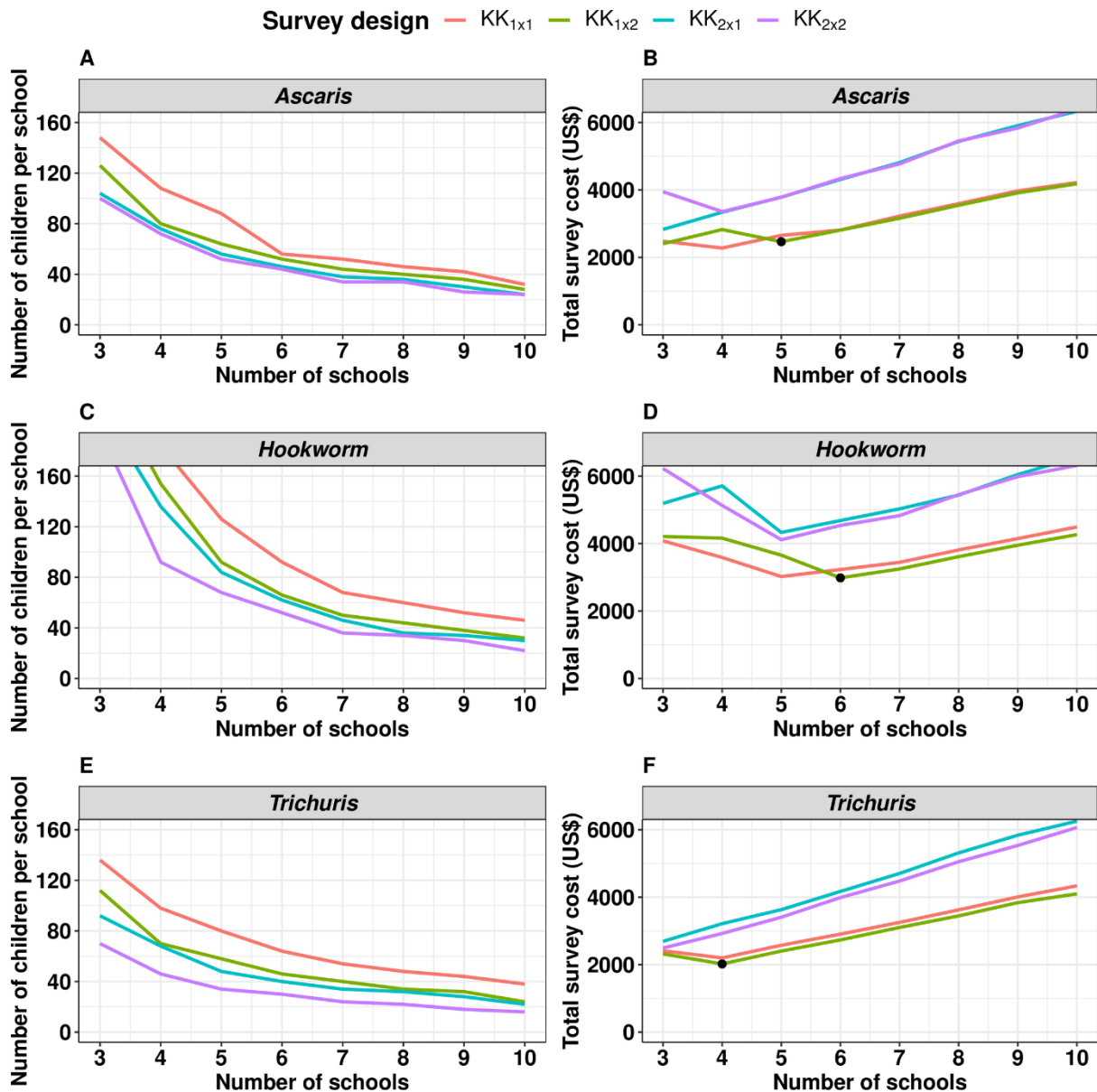
241

242 **Cost-efficient survey design for stopping PC**

243 **Fig 1** illustrates the required number of children per school ($n_{children}$) and the total
244 survey cost (C_{tot}) as a function of the number of sampled schools ($n_{schools}$) to reliably
245 stop PC for the different survey designs for each of the three STH species separately.
246 The panels on $n_{children}$ (**Fig 1A, 1C and 1E**) highlight three important aspects. First,
247 fewer children per school are required when more schools are included in the survey.
248 Second, preparing more smears per sample and collecting more stools per child
249 further reduces the number of children per school. For example, when making program
250 decisions for *Ascaris* infections and sampling 5 schools (**Fig 1A**), 88 children per
251 school are required for a $KK_{1 \times 1}$ survey design, while this number reduces to 64, 56,
252 and 52 when deploying $KK_{1 \times 2}$, $KK_{2 \times 1}$ and $KK_{2 \times 2}$ survey designs, respectively. Third,
253 the minimum required size of the surveys varies across the three STH species. For
254 instance, when sampling 5 schools and deploying a $KK_{1 \times 1}$ survey design, 88 children
255 per school are required for *Ascaris* infections (**Fig 1A**), while this is 126 for hookworm
256 infections (**Fig 1C**) and 80 for *Trichuris* (**Fig 1C**).

257

258 When focusing on the panels representing the total cost C_{tot} (**Fig 1B, 1D and 1F**), two
259 additional aspects become apparent. Survey designs involving a second stool ($KK_{2 \times 1}$
260 and $KK_{2 \times 2}$) are more financially demanding, even though fewer children are required
261 to be sampled. Re-taking the aforementioned example (making program decisions for
262 *Ascaris* infections and sampling 5 schools), the total survey costs for $KK_{1 \times 1}$ survey
263 design equals 2,653 US\$ ($n_{children} = 88$), while survey costs are up to 3,781 US\$ when
264 deploying a $KK_{2 \times 2}$ survey design ($n_{children} = 52$). Second, after an initial drop, the total
265 survey costs further increase as a function of number of schools (for $KK_{1 \times 1}$ and K
266 $K_{1 \times 2}$ survey designs across all STH species). In other words, there is an optimal
267 number of schools that minimizes the costs while ensuring reliable decision-making.
268
269 Combining all these aspects and considering that some schools do not have more
270 than 100 children, $KK_{1 \times 2}$ is the most cost-efficient of all feasible survey designs for all
271 three STH species, but the difference with $KK_{1 \times 1}$ is modest. However, the required
272 sample size ($n_{schools} \times n_{children}$) and the total survey costs varies across the 3 STH
273 species (*Ascaris*: 5 x 64, 2,465 US\$ (**Fig 1B**); hookworm: 6 x 66, 2,983 US\$ (**Fig 1D**);
274 *Trichuris*: 4 x 70, 2,022 US\$ (**Fig 1F**)).



275

276

277 **Fig 1. The identification of the most cost-efficient survey design to reliably stop preventive**

278 **chemotherapy.** This figure illustrates the required number of children per school ($n_{children}$; **Panels A,**

279 **C and E)** and the corresponding total survey cost (C_{tot} ; **Panels B, D and F)** as a function of the number

280 of sampled schools ($n_{schools}$) for different survey designs and soil-transmitted helminths (*Ascaris*:

281 **Panels A and B; hookworm: Panels C and D; Trichuris: Panels E and F).** The survey designs ($KK_{a \times b}$)

282 varied in the number of stool samples per child ($= a$) and the number of Kato-Katz thick smears per

283 sample ($= b$). The black bullet point in Panels B, D and F indicates the most cost-efficient survey design.

284 In other words, the number of schools that minimizes the costs while ensuring reliable decision-making.

285 Note that we assumed that the maximum number of children per school could not exceed 100.

286

287 In **Table 3**, we explored the impact of assumptions about inter-individual variation in
288 EPG and the specificity of KK on what is the optimal survey design ($n_{schools} \times n_{children}$)
289 for the different STH species separately. This table indicates three important patterns
290 when considering *Ascaris*. First, compared to an aggregation of infections within
291 children (k_k), varying as a function of school mean EPG (5 x 64), fixing k_k (= 1.75)
292 resulted in a smaller sample size (3 x 90) and a reduction of the total survey costs of
293 11.5% (= 2,181 US\$/ 2,465 US\$). Second, when assuming a post-control k_k (= 2/3 x
294 varying k_k) resulted in slight increase in both sample size (5 x 68) and total survey
295 costs (1.7% = 2,507 US\$/ 2,465 US\$). Finally, a reduced clinical specificity (= 99.5%)
296 increased the sample size from 5 x 64 to 7 x 62, requiring 38.8% more funds. These
297 patterns were very similar for the other species (**Table 3**).

298

299 We also verified the additional survey costs to further minimize the probability of over-
300 and undertreatment (**Table 3**). Obviously, further minimizing the risk of undertreatment
301 ($E_{undertreat}$) and overtreatment ($E_{overtreat}$) increases the sample size and hence the
302 total survey cost, but not extremely. When we reduce the maximum allowed risk of
303 undertreatment from 0.05 to 0.025, the total survey costs increase by 24% (3,058 US\$
304 vs. 2,465 US\$). When we also reduce the risk of undertreatment from 0.25 to 0.10,
305 the sample size further increases to 78.5% (4,400 US\$ vs. 2,465 US\$). Again, we
306 observed similar patterns for the other two species (**Table 3**).

307

308 **Table 3. The impact of alternative model assumptions on both the required sample size and the total survey costs to reliably stop preventive**
309 **chemotherapy.** This table presents the sample size ($n_{schools} \times n_{children}$) the decision cut-off c and the associated total survey costs (C_{tot}) for different scenarios
310 of parameterizing the aggregation of infections within children (k_k), the clinical specificity of Kato-Katz thick smear and the maximum allowed risk of incorrect
311 decision-making for the different soil-transmitted helminths species separately. As a reference we assumed that k_k varies as a function of school mean eggs
312 per gram of stool (EPG) and that Kato-Katz thick smears has a perfect specificity. We allowed for a risk of undertreatment ($E_{undertreat}$) equal to 0.05 and risk of
313 overtreatment ($E_{overtreat}$) equal to 0.25. Also, the fixed aggregation parameter (Fixed k_k), which captures the variation between individuals within schools, was
314 0.326 for *Ascaris*, 0.257 for hookworm, and 0.532 for *Trichuris*. A lower value of k_k (Lower k_k) means higher inter-individual variation in EPG values, assumed
315 to be 2/3 of k_k in the main analysis.

316

Parameters	<i>Ascaris</i>			Hookworm			<i>Trichuris</i>		
	Sample size	c	C_{tot} (US\$)	Sample size	c	C_{tot} (US\$)	Sample size	c	C_{tot} (US\$)
Aggregation of infections within children (k_k)									
Varying k_k as a linear function of school mean EPG	5 x 64	7	2,465	6 x 66	9	2,983	4 x 70	7	2,022
Fixed k_k	3 x 90	7	2,181	6 x 58	9	2,883	3 x 72	6	1,529
Lower k_k	5 x 68	7	2,507	6 x 70	9	3,033	5 x 58	7	2,403

Parameters	<i>Ascaris</i>			Hookworm			<i>Trichuris</i>		
	Sample size	<i>c</i>	C_{tot} (US\$)	Sample size	<i>c</i>	C_{tot} (US\$)	Sample size	<i>c</i>	C_{tot} (US\$)
Specificity of Kato-Katz thick smear									
100%	5 x 64	7	2,465	6 x 66	9	2,983	4 x 70	7	2,022
99.5%	7 x 62	12	3,422	8 x 70	17	4,044	6 x 58	11	2,883
Risks of incorrect decision-making									
$E_{undertreat} = 0.05; E_{overtreat} = 0.25$	5 x 64	7	2,465	6 x 66	9	2,983	4 x 70	7	2,022
$E_{undertreat} = 0.025; E_{overtreat} = 0.25$	6 x 72	9	3,058	8 x 56	10	3,811	5 x 68	8	2,507
$E_{undertreat} = 0.025; E_{overtreat} = 0.10$	9 x 62	13	4,400	10 x 72	18	5,097	8 x 62	13	3,911

318 **Cost-efficient survey design to declare elimination of STH as a public**
319 **health problem**

320 After having assessed study designs and decision cut-offs for stopping PC, we now
321 turn to study designs for declaring elimination of STH as a public health problem
322 (Table 4; Supplement S4, Figure S1). Compared to the most cost-effective survey
323 designs for stopping PC, two important differences can be noted. First, the $KK_{1 \times 1}$
324 rather than $KK_{1 \times 2}$ survey design is the most cost-efficient choice. Second, the
325 required sample size ($n_{schools} \times n_{children}$) is substantially larger (*Ascaris*: 11 x 84 vs. 5
326 x 64; hookworm: 25 x 52 vs. 6 x 66; *Trichuris*: 19 x 68 vs. 4 x 70). Compared to the
327 total survey costs to reliably stop PC, the cost of surveys to reliably declare EPHP are
328 2.3 (*Ascaris*) to 4.6 (*Trichuris*) times higher.

329

330 **Table 4. The cost-efficient survey design to declare elimination of soil-transmitted helminths as**
331 **a public health problem.** This table represents the required sample size ($n_{schools} \times n_{children}$), the
332 decision cut-off c and the total survey costs (C_{tot}) for four different survey designs ($KK_{a \times b}$) that vary in
333 the number of stool samples per child ($= a$) and the number of Kato-Katz thick smears per sample ($= b$)
334).

STH	Survey design	$n_{schools} \times n_{children}$	c	C_{tot} (US\$)
<i>Ascaris</i>	$KK_{1 \times 1}$	11 x 84	13	5,775
Hookworm	$KK_{1 \times 1}$	25 x 52	17	11,522
<i>Trichuris</i>	$KK_{1 \times 1}$	19 x 68	16	9,346

335

336

337

338

339 We also explored the impact of the assumptions made on the parameterization of the
340 aggregation in infections between children (k_k), the clinical specificity and the allowed
341 risk for both over- and undertreatment on the sample size ($n_{schools} \times n_{children}$) and the
342 total survey cost. Not unexpectedly, we observed similar trends as for reliably stopping
343 PC, and hence we refer to **Supplement S4, Table S1** for more details.

344

345

346 **Discussion**

347

348 We have demonstrated that the current WHO guidelines to M&E STH control
349 programs may not guarantee reliable decision-making for all program targets and STH
350 species. Based on our results, we rather recommend deploying a duplicate KK on one
351 stool when making decisions to stop PC, deploying a single KK on one stool sample
352 when the aim is to verify whether STH has been eliminated as a public health problem.
353 This difference in survey design is in line with other researchers' findings [10] and can
354 be explained by the fact that examining multiple KK smears increases the sensitivity
355 [5,10,17,18] but generally does not affect estimates of the mean intensity of infection
356 [5,10,17,19]. Also, collecting samples on two consecutive days ($KK_{2 \times 1}$ and $KK_{2 \times 2}$)
357 was not cost-efficient, because it requires more funds. In addition, as the required
358 sample size varies across STH species (due to differences in fecundity and
359 aggregation of infections; see **Table 1**), it will be important to ensure sufficient number
360 of schools and children per school to arrive at reliable decision-making for all STH
361 present in the surveyed implementation unit. In other words, the sample size is
362 dictated by the STH species that requires the largest number of schools and children.

363 This means that in an area where all three STH are present, the sample size to reliably
364 stop PC is dictated by hookworms (6 schools and 66 children per school; see **Table**
365 **3**). Note that STH-specific decision cut-offs vary across sample size, and hence they
366 will need to be adapted accordingly. In **Supplement S4, Tables S2 and S3**, we
367 therefore provide the required sample size and the corresponding STH-specific
368 decision cut-offs for areas where more than STH species are considered for decision-
369 making.

370

371 We showed that the aggregation of infections within children (k_k) varies as a function
372 of school mean EPG (see **Supplement S1**), which is in agreement with a recent
373 analysis of the TUMIKIA data. Indeed, this analysis showed that if not considered
374 properly, it leads to poor decision-making due to smaller-than-required sample sizes
375 [20]. Furthermore, the impact of reduced clinical specificity (resulting in increased
376 sample size and total survey cost) was in line with previous work [4], highlighting the
377 need for highly specific diagnostic methods when targeting low-intensity infections
378 settings [21–23].

379

380 Minimizing the risk of making incorrect program decisions (risk of overtreatment and
381 undertreatment) is more financially demanding. However, by improving decision-
382 making, one will also save substantial costs due to unnecessarily distributing tablets
383 or re-establishing the PC program. Furthermore, future research using transmission
384 models is needed to assess the appropriateness of the WHO-recommended
385 thresholds for STH policy.

386

387 The LQAS approach in this simulation study has also been used for population-based
388 decision-making in schistosomiasis [24] and trachoma [25] control programs.
389 Conversely, another approach is the model-based geostatistical methods for M&E of
390 STH, and other NTDs [26,27]. Yet, compared to the LQAS method, this alternative
391 approach adoption required additional expertise in geostatistical analysis [4,27].

392

393 **Conclusion**

394 We confirm that KK is valuable for M&E of STH in low-endemicity infections settings,
395 but the current WHO-recommended survey design may not guarantee reliable
396 decision-making to stop PC, and especially not for declaring EPHP. We, therefore,
397 provide cost-efficient alternative survey designs.

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412 **References**

413

- 414 1. World Health Organization. Ending the neglect to attain the Sustainable
415 Development Goals: A road map for neglected tropical diseases 2021–2030. In:
416 World Health Organization [Internet]. 2021 [cited 10 Feb 2022]. Available:
417 <https://www.who.int/publications/i/item/9789240010352>
- 418 2. World Health Organization. 2030 targets for soil-transmitted helminthiasis
419 control programmes. World Health Organization; 2020 [cited 15 Mar 2022].
420 Available: <https://apps.who.int/iris/handle/10665/330611>
- 421 3. Coffeng LE, Vlaminck J, Cools P, Denwood M, Albonico M, Ame SM, et al. A
422 general framework to support cost-efficient fecal egg count method and study
423 design choices for large-scale STH deworming programs – monitoring of
424 therapeutic drug efficacy as a case study. medRxiv. 2023;
425 2023.01.06.23284253. doi:10.1101/2023.01.06.23284253
- 426 4. Kazienga A, Coffeng LE, de Vlas SJ, Levecke B. Two-stage lot quality
427 assurance sampling framework for monitoring and evaluation of neglected
428 tropical diseases, allowing for imperfect diagnostics and spatial heterogeneity.
429 PLoS Negl Trop Dis. 2022;16: e0010353.
- 430 5. Cools P, Vlaminck J, Albonico M, Ame S, Ayana M, José Antonio BP, et al.
431 Diagnostic performance of a single and duplicate Kato-Katz, Mini-FLOTAC,
432 FECPAKG2 and qPCR for the detection and quantification of soil-transmitted
433 helminths in three endemic countries. PLoS Negl Trop Dis. 2019;13: e0007446.
- 434 6. Khurana S, Sethi S. Laboratory diagnosis of soil transmitted helminthiasis. Trop
435 Parasitol. 2017;7: 86.
- 436 7. Niguse AF, Hailu T, Alemu M, Nibret E, Amor A, Munshea A. Evaluating the

- 437 Performance of Diagnostic methods for soil transmitted helminths in the Amhara
438 National Regional State, Northwest Ethiopia. BMC Infect Dis. 2020.
- 439 8. Dunn JC, Papaiakovou M, Han KT, Chooneea D, Bettis AA, Wyine NY, et al.
440 The increased sensitivity of qPCR in comparison to Kato-Katz is required for the
441 accurate assessment of the prevalence of soil-transmitted helminth infection in
442 settings that have received multiple rounds of mass drug administration. Parasit
443 Vectors. 2020;13: 1–11.
- 444 9. De Vlas SJ, Gryseels B. Underestimation of *Schistosoma mansoni* prevalences.
445 Parasitol Today. 1992;8: 274–277.
- 446 10. Coffeng LE, Malizia V, Vegvari C, Cools P, Halliday KE, Levecke B, et al. Impact
447 of different sampling schemes for decision making in soil-transmitted
448 helminthiasis control programs. J Infect Dis. 2020;221: S531–S538.
- 449 11. Coffeng LE, Levecke B, Hattendorf J, Walker M, Denwood MJ. Survey design
450 to monitor drug efficacy for the control of soil-transmitted helminthiasis and
451 schistosomiasis. Clin Infect Dis. 2021;72: S195–S202.
- 452 12. Leta GT, Mekete K, Wuletaw Y, Gebretsadik A, Sime H, Mekasha S, et al.
453 National mapping of soil-transmitted helminth and schistosome infections in
454 Ethiopia. Parasit Vectors. 2020;13: 1–13.
- 455 13. Denwood MJ, Love S, Innocent GT, Matthews L, McKendrick IJ, Hillary N, et al.
456 Quantifying the sources of variability in equine faecal egg counts: implications
457 for improving the utility of the method. Vet Parasitol. 2012;188: 120–126.
- 458 14. Speich B, Ali SM, Ame SM, Albonico M, Utzinger J, Keiser J. Quality control in
459 the diagnosis of *Trichuris trichiura* and *Ascaris lumbricoides* using the Kato-Katz
460 technique: experience from three randomised controlled trials. Parasit Vectors.
461 2015;8: 1–8.

- 462 15. Knopp S, Salim N, Schindler T, Voules DAK, Rothen J, Lweno O, et al.
463 Diagnostic accuracy of Kato–Katz, FLOTAC, Baermann, and PCR methods for
464 the detection of light-intensity hookworm and *Strongyloides stercoralis* infections
465 in Tanzania. *Am J Trop Med Hyg.* 2014;90: 535.
- 466 16. Levecke B, Coffeng LE, Hanna C, Pullan RL, Gass KM. Assessment of the
467 required performance and the development of corresponding program decision
468 rules for neglected tropical diseases diagnostic tests: Monitoring and evaluation
469 of soil-transmitted helminthiasis control programs as a case study. *PLoS Negl*
470 *Trop Dis.* 2021;15: e0009740.
- 471 17. Moser W, Bärenbold O, Mirams GJ, Cools P, Vlamincck J, Ali SM, et al.
472 Diagnostic comparison between FECPAKG2 and the Kato-Katz method for
473 analyzing soil-transmitted helminth eggs in stool. *PLoS Negl Trop Dis.* 2018;12:
474 e0006562.
- 475 18. Nikolay B, Brooker SJ, Pullan RL. Sensitivity of diagnostic tests for human soil-
476 transmitted helminth infections: a meta-analysis in the absence of a true gold
477 standard. *Int J Parasitol.* 2014;44: 765–774.
- 478 19. Levecke B, Brooker SJ, Knopp S, Steinmann P, Sousa-Figueiredo JC, Stothard
479 JR, et al. Effect of sampling and diagnostic effort on the assessment of
480 schistosomiasis and soil-transmitted helminthiasis and drug efficacy: a meta-
481 analysis of six drug efficacy trials and one epidemiological survey. *Parasitology.*
482 2014;141: 1826–1840.
- 483 20. Truscott JE, Ower AK, Werkman M, Halliday K, Oswald WE, Gichuki PM, et al.
484 Heterogeneity in transmission parameters of hookworm infection within the
485 baseline data from the TUMIKIA study in Kenya. *Parasit Vectors.* 2019;12: 1–
486 13.

- 487 21. Diagnostic target product profiles for monitoring, evaluation and surveillance of
488 schistosomiasis control programmes. [cited 11 Oct 2021]. Available:
489 <https://www.who.int/publications/i/item/9789240031104>
- 490 22. Stuyver LJ, Levecke B. The role of diagnostic technologies to measure progress
491 toward WHO 2030 targets for soil-transmitted helminth control programs. *PLoS*
492 *Negl Trop Dis.* 2021;15: e0009422.
- 493 23. Lim MD, Brooker SJ, Belizario Jr VY, Gay-Andrieu F, Gilleard J, Levecke B, et
494 al. Diagnostic tools for soil-transmitted helminths control and elimination
495 programs: A pathway for diagnostic product development. Public Library of
496 Science San Francisco, CA USA; 2018.
- 497 24. Olives C, Valadez JJ, Brooker SJ, Pagano M. Multiple category-lot quality
498 assurance sampling: a new classification system with application to
499 schistosomiasis control. *PLoS Negl Trop Dis.* 2012;6: e1806.
- 500 25. Myatt M, Mai NP, Quynh NQ, Nga NH, Tai HH, Long NH, et al. Using lot quality-
501 assurance sampling and area sampling to identify priority areas for trachoma
502 control: Viet Nam. *Bull World Health Organ.* 2005;83: 756–763.
- 503 26. Diggle PJ, Amoah B, Fronterre C, Giorgi E, Johnson O. Rethinking neglected
504 tropical disease prevalence survey design and analysis: a geospatial paradigm.
505 *Trans R Soc Trop Med Hyg.* 2021;115: 208–210.
- 506 27. Johnson O, Fronterre C, Amoah B, Montresor A, Giorgi E, Midzi N, et al. Model-
507 based geostatistical methods enable efficient design and analysis of prevalence
508 surveys for soil-transmitted helminth infection and other neglected tropical
509 diseases. *Clin Infect Dis.* 2021;72: S172–S179.
- 510
511

512 **Funding**

513 Bruno Levecke acknowledges funding from Ghent University starting grant
514 (www.ugent.be) and Luc E. Coffeng from the Dutch Research Council (NWO, grant
515 016.Veni.178.023). In addition, the financial support for the dataset used in this study
516 was provided by Schistosomiasis Control Initiative, the Partnership for Child
517 Development, the Children's Investment Fund Foundation, the End Neglected Tropical
518 Diseases Fund, and UKAID-DFID.

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538 **Supporting Information Legends**

539

540 **Supplementary information S1.** Quantification of between and within -school
541 variation in faecal egg counts

542

543 **Supplementary information S2.** The mathematical backbone of the simulation
544 framework

545

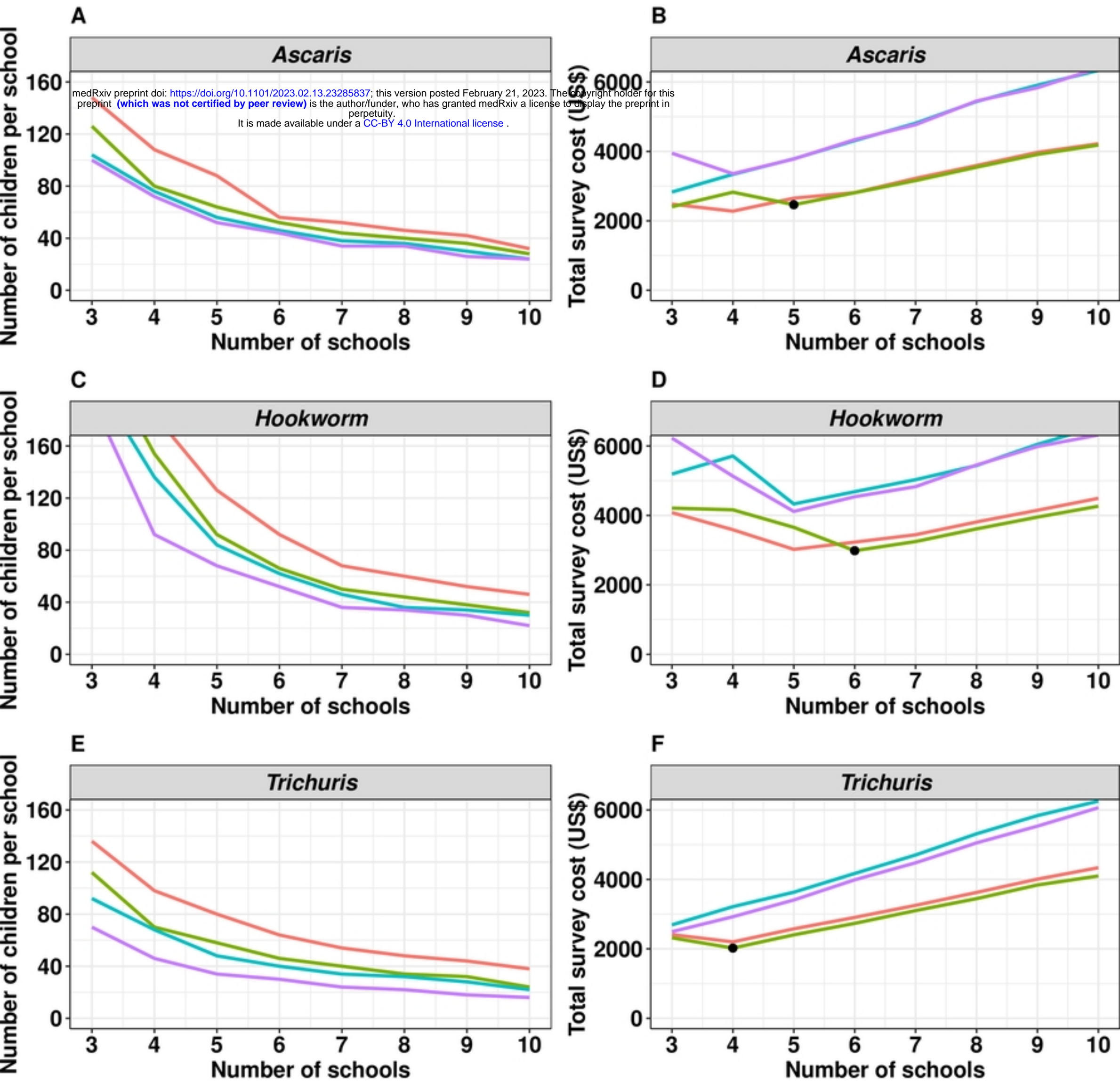
546 **Supplementary information S3.** Estimation of total survey cost to monitor and
547 evaluate STH when using Kato-Katz thick smear

548

549 **Supplementary information S4.** Additional tables and figures

550

Survey design — KK_{1x1} — KK_{1x2} — KK_{2x1} — KK_{2x2}



Figure