CHAPTER FIVE

Mathematical Inference on Helminth Egg Counts in Stool and Its Applications in Mass Drug Administration Programmes to Control Soil-Transmitted Helminthiasis in Public Health

Bruno Levecke*,[1], Roy M. Anderson[§], Dirk Berkvens[¶], Johannes Charlier*, Brecht Devleesschauwer*,[||], Niko Speybroeck[||], Jozef Vercruysse*, Stefan Van Aelst[#],[**]

*Department of Virology, Parasitology and Immunology, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium
[§]Imperial College London, Department of Infectious Disease Epidemiology, Faculty of Medicine, School of Public Health St Mary's Campus, Norfolk Place, London, UK
[¶]Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium
[||]Institute of Health and Society (IRSS), Université catholique de Louvain, Brussels, Belgium
[#]Department of Applied Mathematics, Computer Science and Statistics, Faculty of Sciences, Ghent University, Ghent, Belgium
**Department of Mathematics, Faculty of Sciences, KU Leuven University, Leuven, Belgium
[1]Corresponding author: E-mail: bruno.levecke@ugent.be

Contents

Advances in Parasitology, Volume 87
ISSN 0065-308X
http://dx.doi.org/10.1016/bs.apar.2015.01.001

Abstract

In the present study, we present a hierarchical model based on faecal egg counts (FECs; expressed in eggs per 1 g of stool) in which we first describe the variation in FECs between individuals in a particular population, followed by describing the variance due to counting eggs under a microscope separately for each stool sample. From this general framework, we discuss how to calculate a sample size for assessing a population mean FEC and the impact of an intervention, measured as reduction in FECs, for any scenario of soil-transmitted helminth (STH) epidemiology (the intensity and aggregation of FECs within a population) and diagnostic strategy (amount of stool examined (~sensitivity of the diagnostic technique) and examination of individual/pooled stool samples) and on how to estimate prevalence of STH in the absence of a gold standard. To give these applications the most wide relevance as possible, we illustrate each of them with hypothetical examples.

1. INTRODUCTION

The soil-transmitted helminths (STHs, *Ascaris lumbricoides*, *Trichuris trichiura* and the two hookworm species, *Necator americanus* and *Ancylostoma duodenale*) cause the highest burden among all neglected tropical diseases

(NTDs) (Murray et al., 2012). Recent global numbers indicate that in 2010 more than 1.4 billion people were infected with at least one of the four STH species, resulting in a global burden of approximately 5.2 million disability-adjusted life years (DALYs) (∼20% of total DALYs attributable to NTDs) (Pullan et al., 2014). Mass drug administration (MDA) programmes in which a single oral dose of albendazole (ABZ) or mebendazole (MBZ) — the drugs of choice for STHs — are periodically administered to pre-school and school-aged children are the main strategy to control the morbidity caused by STHs (WHO, 2011), and these programmes have recently received increased political and scientific attention. The World Health Organization (WHO) has devised a roadmap to guide implementation of the policies and strategies set out in a global plan to combat NTDs (period 2008–2015), and more than 70 pharmaceutical companies, governments and global health organizations committed to supporting this roadmap in the London Declaration on NTDs in January 2012 by sustaining or expanding drug donation programmes (WHO, 2012a; NTD Partner Website, 2013). With this growing attention, WHO aims to increase the coverage of the pre-school and school-aged children in need of drug administration from 30% (estimated coverage in 2010; WHO, 2012b) to at least 75% by 2020, and to ultimately eliminate soil-transmitted helminthiasis as a public health problem in children (WHO, 2012c). These pledges of drug donations are at place, but this global upscale of MDA programmes also creates the need for a monitoring system that allows programme managers, policymakers and donors of the drugs to assess whether the objectives are being met and, if necessary, to adjust the implemented strategy (WHO, 2011). Thus, it will be imperative to periodically assess STH infections by means of prevalence and infection intensity to determine whether the MDA programme progresses as anticipated.

MDA programmes are currently poorly monitored, and one of the main reasons for this lack of monitoring systems is the absence of a framework that guides healthcare decision-makers in designing surveys. Development of such a framework, however, is not straightforward. First, due to the heterogeneity of STH infections, it will be impossible to apply one survey design to all implementation units of MDA, both within and between countries (Brooker et al., 2010; Hürlimann et al., 2011; Pullan and Brooker, 2012; Pullan et al., 2014). Moreover, as discussed by Bergquist et al. (2009), survey designs applied in an early stage of MDA programmes may not guarantee a reliable assessment of STH infections in a later stage, as both prevalence and intensity of infections will decrease over consecutive rounds of MDA.

Second, demonstration and quantification of eggs in stool is the current standard means to diagnose STH infections; however, this diagnosis is thwarted by the absence of a gold-standard technique, day-to-day variation in egg excretion and heterogenous distribution of eggs within stool samples (Sinniah, 1982; Ye et al., 1997; Krauth et al., 2012). Finally, MDA programmes typically operate in resource-constrained settings, and hence it is indispensable that healthcare decision-makers have some pliancy to minimize both financial and technical resources, while assuring a reliable assessment of the progress made.

The overall aim of this chapter is to develop a mathematical framework based on helminth egg counts in stool that allows healthcare decision-makers to adapt their survey design according to both local STH epidemiology and resources. Specifically, we will first list the most important sources of variability in egg counts in stool, and how they affect the design of studies. Next, we will outline a general mathematical framework for helminth egg counts in stool. From this general framework, we will continue by working out a selected number of applications for surveys designed to monitor MDA programmes to control STHs. To give these applications the most wide relevance as possible, we will illustrate each of them with hypothetical examples.

2. SOURCES OF VARIABILITY IN EGG COUNTS

In principle, the presence of STH eggs in stool is the result of at least one adult female in the gastrointestinal tract that is laying eggs. However, the number of eggs that is excreted in stool and which is eventually counted under the microscope is affected by a variety of factors. These can be classified in two groups of variation, including variation due to the egg excretion and due to the egg counting procedure. The variation in egg excretion is mainly due to biological factors, whereas the variation in the egg counting procedure is due to technical factors. We will address both sources of variability separately.

2.1 Sources of variability in egg excretion

Important sources that affect the number of eggs excreted in stool include fecundity of adult female worms and host-parasite-environment interactions. We will illustrate each of the two aspects by field data obtained during an epidemiological survey conducted in three countries in East

Africa (Ethiopia, Kenya and Uganda). The main objective of this survey was to investigate the distribution and heterogeneity of co-infection with *Plasmodium falciparum* and helminths, including STHs and *Schistosoma mansoni*. The presence and intensity of these helminth infections were determined by examination of one stool sample per child with duplicate Kato-Katz thick smears. Brooker et al. (2012) describe this epidemiological survey more in detail. We will use data obtained in Kenya which are made publicly available at http://www.thiswormyworld.org and are composed of 17,871 children across 178 schools (median number of children per school = 104).

There is a manifest difference in fecundity of adult female worms between the different STH species. An adult female *A. lumbricoides* worm produces approximately 200,000 eggs per day, whereas this ranges from 25,000 to 30,000, and from 9,000 to 10,000 eggs per day for *A. duodenale* and *N. americanus*, respectively. The daily egg output is the lowest for *T. trichiura*, with a daily egg output ranging from 3,000 to 5,000 eggs (Bethony et al., 2006). These differences in fecundity translate into differences in mean faecal egg counts (FECs) per gram of stool (EPG) across the STH species. For example, the mean FECs for *A. lumbricoides* in the Kenyan survey equalled 5,672 EPG ($N = 582$), whereas this was 332 EPG ($N = 2,086$) and 241 ($N = 1,672$) for hookworm and *T. trichiura*, respectively. As a consequence of this difference in fecundity, a survey designed for assessing *A. lumbricodes*, may not always allow assessing *T. trichiura* or hookworm infections with an equal level of precision (Levecke et al., 2011a). This is challenging, particularly because mixed STH infections are very common. In the Kenyan survey, 20% of the subjects infected with STH were excreting eggs of more than one STH species.

Although it is difficult to unravel the contribution of host, parasite and environment factors to the variation in egg excretion separately, their impact on egg excretion is pertinent. The most important consequences of this complex interplay of host-parasite-environment are a heterogenous distribution of FECs both within and between populations (e.g. school and community), a day-to-day variation in egg excretion, and a heterogeneous distribution in stool. Generally, a minority of the individuals excrete the majority of the eggs, and as illustrated in Figure 1 for school ID1175 of the survey in Kenya; this typically results in a skewed distribution of FECs. In this school 20% of children contributed to the total number of eggs excreted for *A. lumbricoides* and hookworm, and 96.7% for *T. trichiura*. This skewed distribution of FECs expressed in EPG can be modelled using a negative binomial distribution or

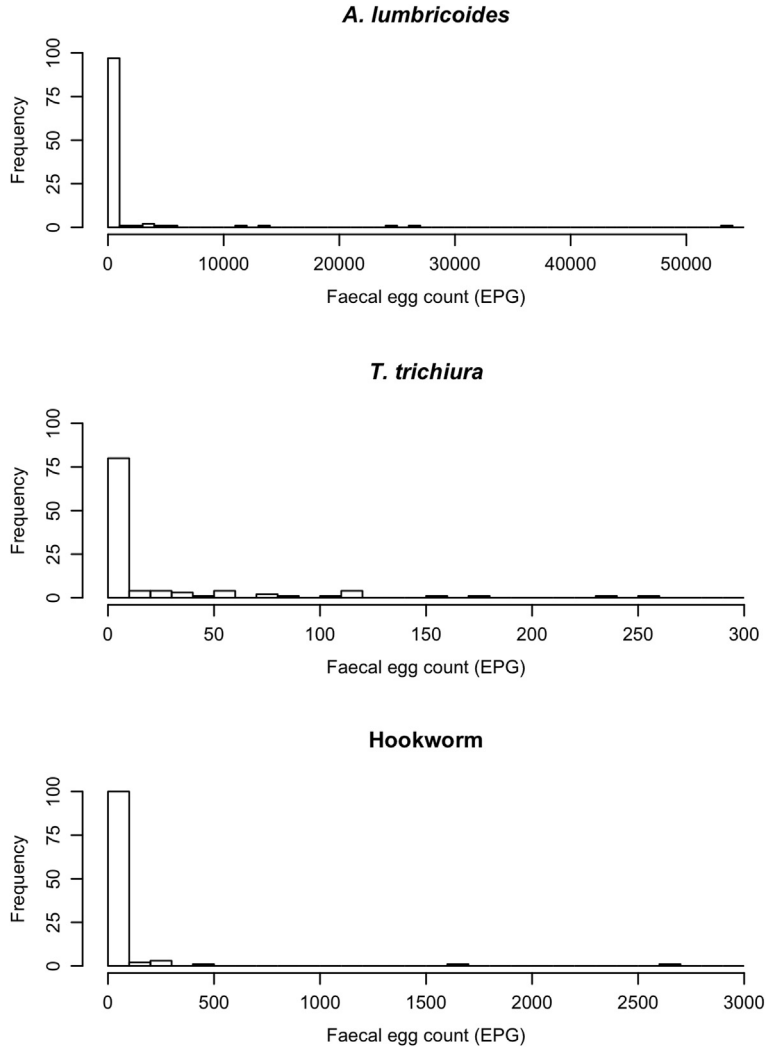

Figure 1 Histogram of individual faecal egg counts (expressed in eggs per gram of stool (EPG)) of *Ascaris lumbricoides*, *Trichuris trichiura* and hookworm of a random sample of 108 subjects from school ID1175 included in the epidemiological survey conducted in Kenya (Brooker et al., 2012). Note the differential skewness across the three soil-transmitted helminth species.

zero-inflated count distributions. In our framework, we will assume that FECs within a population follow a negative binomial distribution, which may be defined by a mean population FEC μ and an aggregation parameter k (Dobson et al., 2009; Torgerson et al., 2005). The parameter k can be

calculated as a function of the mean and variation of the individual FECs through the moment estimator ($k = \text{mean FEC}^2/(\text{variance FEC} - \text{mean FEC})$), and is inversely correlated with aggregation: low values indicate a high level of aggregation and high values indicate a more random distribution of FECs. In school ID1175, the mean FEC and k equalled 1,400 EPG and 0.05 for *A. lumbricoides*, 20 EPG and 0.18 for *T. trichiura* and 56 EPG and 0.03 for hookworm. However, these values for mean FEC and k vary considerably among populations. The mean FEC across the 178 schools included in the survey in Kenya, ranged from 0 EPG up to 6,173 EPG for *A. lumbricoides*, up to 499 EPG for *T. trichiura* and up to 482 EPG for hookworm infections. As illustrated in Figure 2, infections are also aggregated between schools, the minority of the schools accounting for the majority of the egg excretion (20% of the schools cover 97.2% of total number of *A. lumbricoides* eggs excreted; for *T. trichiura* and hookworm this was 60.9% and 83.0%, respectively). The values for k were up to 0.32 for *A. lumbricoides*, 0.45 for *T. trichiura* and 0.28 for hookworm infections. These values for k increased as a function of increasing mean FEC (Figure 2). These skewed and species-specific FEC distributions have three important implications. First, they highlight that surveys will have to be designed for each STH species and population separately. Second, they pose a serious risk of bias, as the mean of a small subsample of individual FECs is very likely to underestimate the mean population FEC (Gregory and Woolhouse, 1993). Finally, current formulae to calculate sample sizes are based on a normal distribution of the mean FEC (central limit theorem), an approximation which may be very poor when sample size is small, especially in some scenarios of mean population FEC and k (see Figure 3).

The consequences of day-to-day variation in egg excretion and heterogeneous distribution of eggs within stool on FEC results are well known and documented (Sinniah, 1982; Ye et al., 1997; Krauth et al., 2012). To minimize the day-to-day variation in individual FECs, it has been suggested to examine several samples per subject, collected over consecutive days (Booth et al., 2003; Knopp et al., 2008). To minimize the heterogeneous distribution within a stool sample, one can either examine multiple Kato-Katz thick smears per stool sample or use a diagnostic technique that allows examining a larger amount of stool such as the FLOTAC technique (up to 0.5 g) (Knopp et al., 2009; Cringoli et al., 2010). It is expected that these measures will decrease the variation between individual FECs, and hence increase the precision of population

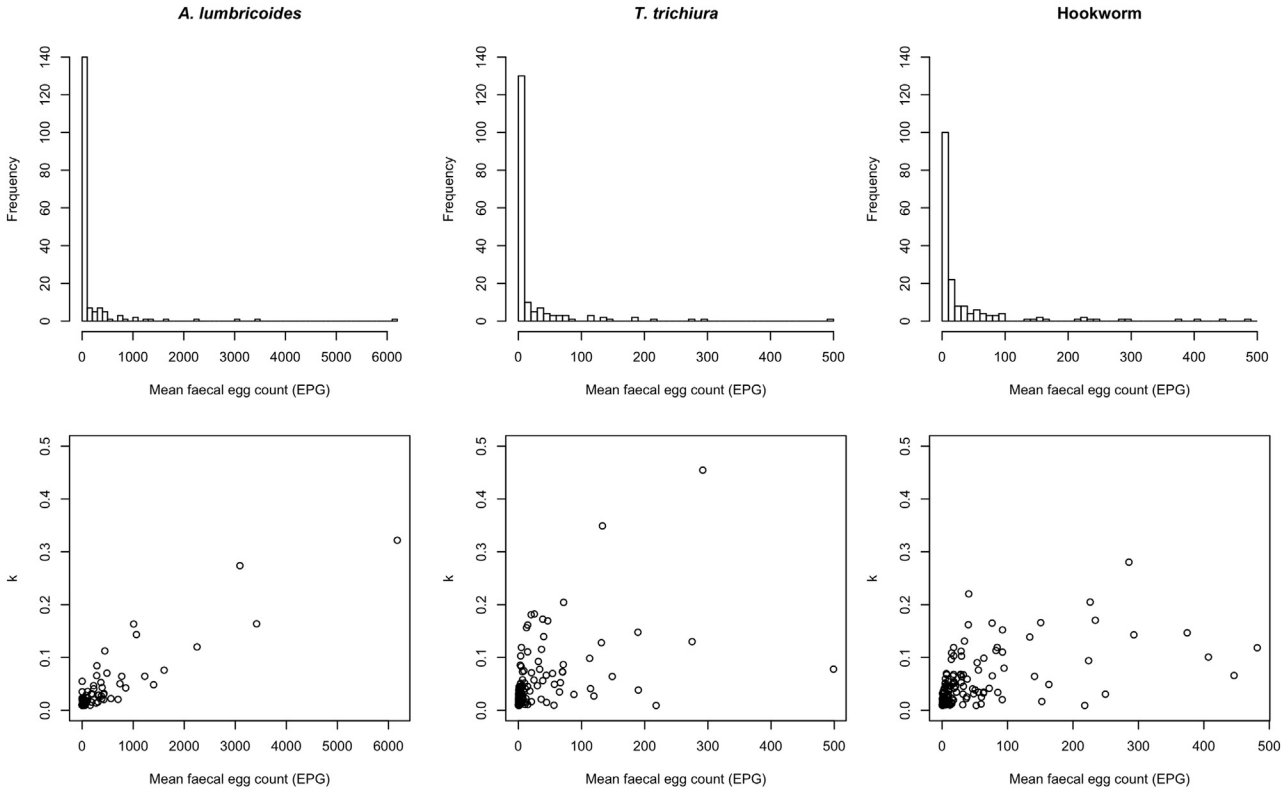

Figure 2 Histogram of mean school faecal egg counts (FECs, expressed in eggs per gram of stool (EPG)) of *Ascaris lumbricoides*, *Trichuris trichiura* and hookworm of all 178 schools included in the epidemiological survey conducted in Kenya (Brooker et al., 2012) and scatter plots of aggregation parameter *k* as a function of mean school FECs (bottom graphs). Note the differential skewness across the three soil-transmitted helminth species.

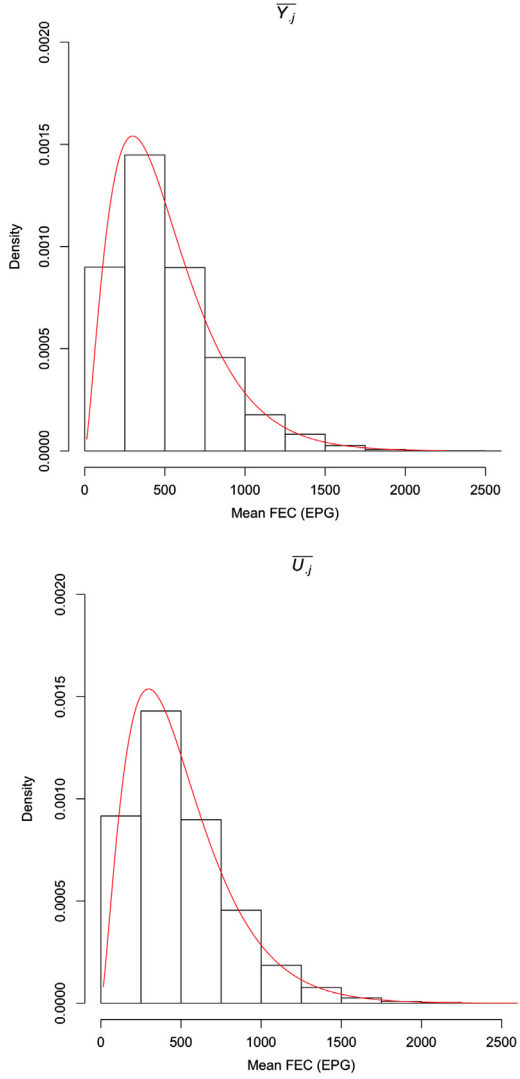

Figure 3 The histograms illustrate the probability density of the mean of observed individual faecal egg counts (FECs; expressed in eggs per g of stool (EPG)) $(= \overline{Y_{.j}})$ and mean of observed pooled FECs $(= \overline{U_{.j}})$ for *Ascaris lumbricoides* generated from 10,000 theoretical surveys in which 50 subjects were sampled from an infinitive population j ($\mu_j = 500$ EPG and $k_j = 0.05$). Both individual $(N_{j_{ind}} = 50)$(top graph) and pooled stool samples $(m_j = 5, N_{j_{pool}} = 10)$ (bottom graph) were examined using a single Kato-Katz thick smear ($f_j = 0.0417$ g). The red (light gray in print versions) lines describe the corresponding probability density of a gamma distribution with $\gamma_{j_{ind}} = \frac{\mu_j \cdot N_{j_{ind}}}{1/f_j+1+\mu_j/k_j}$ and $\theta_{j_{ind}} = \frac{1/f_j+1+\mu_j/k_j}{N_{j_{ind}}}$ (top graph), and $\gamma_{j_{pool}} = \frac{\mu_j \cdot N_{j_{pool}}}{1/f_j+(1+\mu_j/k_j)/m_j}$ and $\theta_{j_{pool}} = \frac{1/f_j+(1+\mu_j/k_j)/m_j}{N_{j_{pool}}}$ (bottom graph).

mean FEC estimates. As consequence of this, a smaller sample size will be required to estimate population mean FEC based on a duplicate Kato-Katz thick smear with an equal level of precision. These measures, however, will have less impact on the accuracy of these estimates, because the variation in egg excretion across days and within stool is random (subjects will not systematically excrete more eggs at the first sample collection and eggs are not systematically found in one particular part of stool). To verify this, we have summarized the mean *A. lumbricoides* FECs and their corresponding variance based on the examination of a single Kato-Katz thick smear and a duplicate Kato-Katz thick smear on the same stool sample in 10 schools from the epidemiological survey in Kenya (Table 1). Compared to a duplicate Kato-Katz thick smear, a single Kato-Katz resulted in a higher variance in eight out of 10 schools. In the two remaining schools (ID1020 and ID1023), a lower variance was found. Mean FEC based on a single and duplicate Kato-Katz thick smear were comparable, and were not systematically under- or overestimated, suggesting that indeed an increase in sampling and diagnostic effort increases the precision of population mean FECs estimates, but not the accuracy. For a more detailed analysis we refer the reader to a recent study by Levecke et al. (2014a), in which a comparison of FECs was made based on collection of one or two stool samples that were processed with single or duplicate Kato-Katz thick smears.

Table 1 The mean and variance in *Ascaris lumbricoides* faecal egg counts (FECs, expressed in eggs per gram of stool (EPG)) in 10 schools for a single and a duplicate Kato-Katz thick smear. The data were collected during an epidemiological survey conducted in Kenya (Brooker et al., 2012)

		Single Kato-Katz		Duplicate Kato-Katz	
School ID	***N***	**Mean FEC (EPG)**	**Variance**	**Mean FEC (EPG)**	**Variance**
1002	102	44	$2.02 \cdot 10^5$	43	$1.88 \cdot 10^5$
1003	100	345	$3.87 \cdot 10^6$	306	$3.30 \cdot 10^6$
1005	100	423	$6.34 \cdot 10^6$	427	$5.75 \cdot 10^6$
1006	100	1	51.8	<1	12.9
1007	104	<1	11.0	1	6.65
1010	108	184	$1.58 \cdot 10^6$	112	$5.37 \cdot 10^5$
1020	107	3	$7.79 \cdot 10^2$	4	$1.47 \cdot 10^3$
1023	104	175	$6.33 \cdot 10^5$	225	$1.08 \cdot 10^6$
1026	103	850	$3.51 \cdot 10^7$	564	$1.44 \cdot 10^7$
1027	103	339	$5.07 \cdot 10^6$	329	$3.75 \cdot 10^6$

2.2 Sources of variability in egg counting procedure

The egg counting procedure provides an estimate of the number of eggs in 1 g of stool by examining a small amount of stool (<1 g). Generally, it consists of three consecutive steps, including (1) preservation of stool, (2) processing the samples with a diagnostic technique that allows counting eggs under the microscope and (3) microscopically counting the eggs. Each of these steps can potentially affect the observed number of eggs excreted per g of stool, resulting in imprecise or inaccurate FEC estimates. Traditionally, samples are prepared fresh within 24 h after stool collection, but when this is not possible samples are often preserved in formalin. Although this allows to postpone the examination of samples, it thwarts the FECs. Indeed, examination of preserved samples resulted in lower FEC compared to FEC obtained after examining fresh samples (Albonico et al., 2013). Currently, a variety of diagnostic techniques have been applied to process human stool (e.g. Kato-Katz thick smear (Katz et al., 1972; WHO, 1991), McMaster (MAFF, 1986; Levecke et al., 2011b), Mini-FLOTAC (Cringoli et al., 2013; Barda et al., 2013) and FLOTAC techniques (Cringoli et al., 2010)), but they rarely result in identical FECs. These techniques differ in many aspects, but we will focus on the five most important differences that contribute to the variation in FECs between these techniques, including mode of action, amount of stool that is sampled, the way this amount of stool is determined, the amount of stool that is examined and the use of a counting apparatus (Table 2).

When using the Kato-Katz technique a simple stool smear is made, whereas for McMaster and (Mini-)FLOTAC stool is suspended in a solution of which the density is higher than that of STH eggs, allowing eggs to float to the surface. Although techniques based on flotation result in more clean microscopic views (debris will not always float), and hence reduce the inter-rater variability in FECs, they may result in lower FECs compared to Kato-Katz thick smear. For example, Levecke et al. (2011b) noticed that unfertilized *A. lumbricoides* eggs did often not float to the surface, contributing to discrepancies in FECs observed between flotation-based techniques and Kato-Katz technique (Knopp et al., 2009). Sampling a larger proportion of the stool will result in more precise estimates of FECs, as it will overcome the heterogeneous distribution of eggs in stool (see also Table 1). Of the four currently applied diagnostic techniques, Kato-Katz processes the least stool ($\sim$0.0417 g), FLOTAC the most (in theory the entire stool sample can be processed). To assess FECs by means of EPG, it is essential that a known mass of stool is processed. For the flotation-based techniques, a

Table 2 Mode of action, amount of stool that is sampled, the way this amount of stool is determined, the amount of stool that is examined and the use of a counting apparatus for four commonly diagnostic techniques used for the detection and quantification of soil-transmitted helminth eggs in human stool

	Kato-Katz	McMaster	Mini-FLOTAC	FLOTAC
Mode of action	Smear	Flotation	Flotation	Flotation
Mass of stool sampled (g)	0.0417	2	2	Entire stool sample
Determination of amount of stool	Volume	Mass	Mass	Mass
Sieving stool prior/after the determination of amount of stool	Prior	After	After	After
Mass of stool examined (g)	0.0417	0.02	0.1	0.5
Multiplication factor (=1/mass of stool examined)	24	50	10	2
Counting apparatus	No	Yes	Yes	Yes

known mass of stool is weighed, after which the stool is suspended in flotation solution and sieved to withhold the large debris. This is in contrast with the Kato-Katz technique. For this technique, the stool samples are first sieved, after which a known volume of sieved stool is sampled. The difference in sequence of sieving, contributes to the higher FECs observed for Kato-Katz technique compared to the flotation-based techniques (Knopp et al., 2009; Levecke et al., 2011b). Measuring stool volumetrically introduces additional variation across FECs, as the density of stool is not fixed across individuals (Levecke et al., 2011b). Examination of a larger proportion of the stool will result in more precise estimates of FECs. Of the four currently applied diagnostic techniques, the least stool is examined with McMaster (= 0.02 g), the most with FLOTAC (up to 0.5 g). Finally, in contrast to the Kato-Katz technique, the flotation-based techniques make use of a counting apparatus in which the number of eggs is counted within grids. These grids are designed to guide the laboratory technicians in examining the samples, and hence minimize the probability of missing eggs or counting eggs twice. Despite these specific measures to increase the agreement in FECs between laboratory technicians and laboratories, variation in FEC due to human error remains considerable (Bogoch et al., 2006; Levecke et al., 2011b). Therefore, it has been recommended to re-examine a subset of the samples by a senior laboratory technician to further minimize the inter-rater variation in FECs.

Currently, we lack detailed studies that can provide estimates of bias and variation introduced by each of the aforementioned aspects across a large number of laboratories for STH diagnosis. Moreover, keeping the final application of the mathematical framework in mind, we would like to minimize the number of parameters that needs to be defined by the end user. In the mathematical framework described in this chapter, we will assume that the number of eggs observed by microscopy follows a Poisson distribution which is defined by a parameter λ. This parameter, however, depends on how much stool is examined, and hence which diagnostic technique is applied. For example, if the true FEC in a stool sample equals 500 EPG, than the expected number of eggs λ to be counted under the microscope in 0.02 g of stool equals 10 (=true FEC·amount of stool examined (in g) = 500·0.02) and 250 when 0.5 g of stool is examined (=500·0.5). To obtain the FECs expressed in EPG, the observed egg counts are subsequently multiplied by 50 and 2 (= 1/amount of stool examined (in g)), respectively. However, due to random variation the FECs obtained after re-processing the same sample will not be identical. This assumption that the number of eggs observed by microscopy follows a Poisson distribution has been mainly studied in veterinary parasitology (Morgan et al., 2005; Torgerson et al., 2012), and as we will illustrate, it already explains most of the variation in FECs reported in field studies. Figure 4 illustrates the variation in FECs obtained by theoretically processing 0.02 (McMaster) and 0.5 g (FLOTAC) of the same stool sample 100 times. To explore the change in variation as a function of FEC, we show the variation in a stool sample in which the true underlying FEC equals 5, 50 and 500 EPG. From this figure, we can deduce that (1) on average the observed FECs approaches the true FEC, (2) the variance ($\sim$1/precision) increases as a function of decreasing amount of stool that is examined and increasing true FECs and (3) the sensitivity increases as a function of increasing amount of stool examined and the true FEC. Each of these aspects has been observed in field data (see Table 1, Levecke et al., 2011b, 2014a), highlighting that the assumption that the number of eggs counted by microscopy follows a Poisson distribution is justified.

3. GENERAL MATHEMATICAL FRAMEWORK FOR FECs

In this section we will describe a general mathematical framework based on FECs. Given that the current diagnostic techniques only

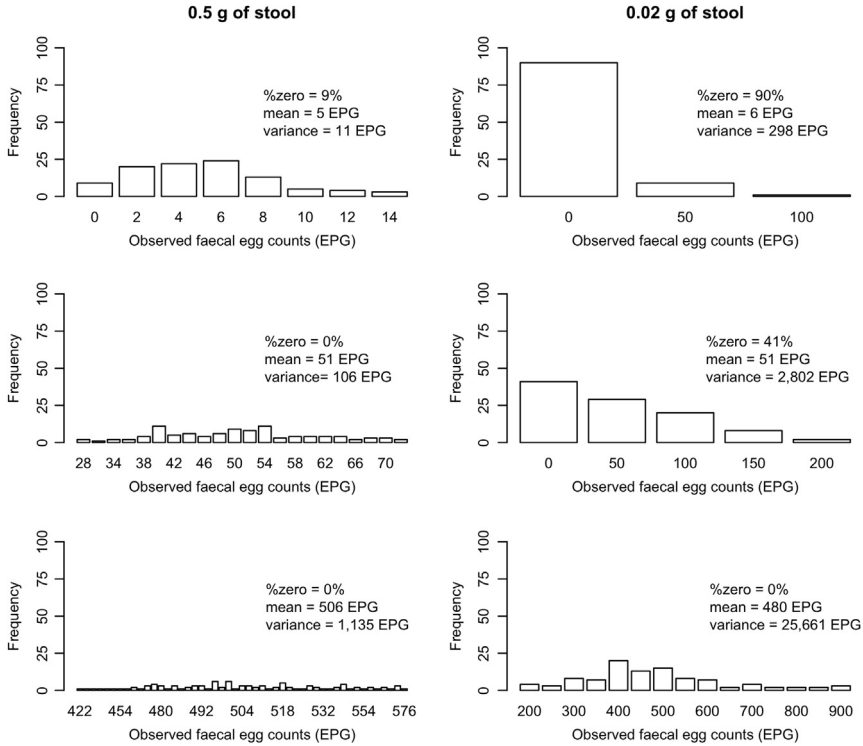

Figure 4 Histogram of 100 theoretical faecal egg counts (FECs, expressed in eggs per gram of stool (EPG)) obtained by examining 0.02 and 0.5 g of stool. It was assumed that the number of eggs observed by microscopy follows a Poisson distribution with λ equal to true FEC · amount of stool examined (in g). The true FEC was set at 5 (top graphs), 50 (middle graphs) and 500 EPG (bottom graphs). We provide the proportion of zero FECs (%*zero*), mean and variance of FECs.

provide estimates of the number of eggs in 1 g of stool as a proxy of egg excretion, we will express these FECs in EPG. We will assume that individual FECs within a population follow a negative binomial distribution, that there is no day-to-day variation in FECs, and that variation in egg counts from the same stool sample under the microscope is completely explained by a Poisson distribution. To give healthcare decision-makers more flexibility to adapt their survey design, the framework will be described for the examination of both individual and pooled stool samples. A list of all variables in this general framework is given in Table 3.

Table 3 The variables included in the hierarchical models describing faecal egg counts (FECs, expressed in eggs per gram of stool (EPG)) prior and after an intervention, and for examination of individual and pooled stool samples

	Definition
e_j	True efficacy of an intervention in a population j
f_j	Amount of stool (in g) collected from a population j that is microscopically examined
μ_j	True population mean FEC (in EPG) in a population j prior to an intervention
λ_{ij}	True number of eggs in f_j g of stool of an individual i of population j prior to an intervention
λ'_{ij}	True number of eggs in f_j g of stool of an individual i of population j after an intervention
λ_{lj}	True number of eggs in f_j g of a pooled stool sample l from a population j prior to an intervention
λ'_{lj}	True number of eggs in f_j g of a pooled stool sample l from a population j after an intervention
k_j	True aggregation of FEC (in EPG) in a population j prior to an intervention
m_j	The number of stool samples collected from individuals of population j that are pooled into one sample
$n_{j_{ind}}$	The number of individuals in a population j
$n_{j_{pool}}$	The number pools in a population j
U_{lj}	Observed FEC (in EPG) obtained by examining f_j g of a pooled stool sample l from a population j prior to an intervention
U'_{lj}	Observed FEC (in EPG) obtained by examining f_j g of a pooled stool sample l from a population j after an intervention
V_{lj}	True FEC (in EPG) in a pooled stool sample l from a population j prior to an intervention
V'_{lj}	True FEC (in EPG) in a pooled stool sample l from a population j after an intervention
X_{ij}	True FEC (in EPG) of an individual i of population j prior to an intervention
X'_{ij}	True FEC (in EPG) of an individual i of population j after an intervention
Y_{ij}	Observed FEC (in EPG) obtained by examining f_j g of stool of an individual i of a population j prior to an intervention
Y'_{ij}	Observed FEC (in EPG) obtained by examining f_j g of stool of an individual i of a population j after an intervention
Z_{ij}	Observed number of eggs counted under the microscope in f_j g of stool from an individual i of a population j prior to an intervention
Z'_{ij}	Observed number of eggs counted under the microscope in f_j g of stool from an individual i of a population j after an intervention
Z_{lj}	Observed number of eggs counted under the microscope in f_j g of a pooled stool sample l from a population j prior to an intervention
Z'_{lj}	Observed number of eggs counted under the microscope in f_j g of a pooled stool sample l from a population j after an intervention

3.1 Faecal egg counts

3.1.1 FECs based on individual stool samples

Generally, the framework describes a hierarchical model based on FECs in which we first describe the distribution of FECs between individuals, followed by describing the process of counting eggs under a microscope separately for each stool sample. For the distribution of FECs between individuals, we assume that FECs X_{ij} (in EPG) between individuals $i = 1, 2, \ldots, n_{j_{ind}}$ of a population j follow a negative binomial distribution with a population mean μ_j and aggregation parameter k_j. For the process of counting eggs, we assume that the number of eggs observed by microscopy in stool of an individual i of population j is Poisson distributed with λ_{ij} as the true underlying number of eggs in f_j g of stool equal to $f_j \cdot X_{ij}$, where f_j is the amount of stool examined (in g) and where X_{ij} is the true FEC expressed in EPG for an individual i. Because f_j is always less than 1 g we multiply these observed egg counts under the microscope Z_{ij} by a factor $1/f_j$ to obtain the observed FEC Y_{ij} in EPG for an individual i in a population j. This results in the following hierarchical model for observed FECs Y_{ij},

$$\begin{aligned} X_{ij} &\sim NB\left(\mu_j,\ k_j\right) \\ Z_{ij}\left|X_{ij}\right. &\sim Poiss\left(\lambda_{ij}\right) \\ &\sim Poiss\left(f_j \cdot X_{ij}\right) \\ Y_{ij} &= \frac{Z_{ij}}{f_j} \end{aligned} \tag{1}$$

It can be derived that the expected value of Y_{ij} equals μ_j, the mean of the negative binomial distribution describing the FECs between individuals in a population j. Due to the additional variation introduced by the egg counting process, the variance of the observed individual FECs Y_{ij} will be larger than those of true individual FECs $X_{ij} (= \mu_j + \mu_j^2/k_j)$. The variance of Y_{ij} given Eqn (1) is described below in Eqn (2). The variance of Y_{ij} will increase with increasing levels of egg excretion (μ_j) and aggregation ($\sim 1/k_j$), and with decreasing quantity of stool examined (f_j). The quantity of stool that is examined can be increased by changing diagnostic technique (Kato-Katz thick smear ($f_j = 0.0417$ g) vs FLOTAC ($f_j = 0.5$ g)) or by repeatedly processing the same sample (single ($f_j = 0.0417$ g) vs duplicate ($f_j = 0.0834$ g) Kato-Katz thick smears).

$$\begin{aligned} E\left[Y_{ij}\right] &= \mu_j \\ Var\left[Y_{ij}\right] &= \mu_j \cdot \left(\frac{1}{f_j} + 1 + \frac{\mu_j}{k_j}\right) \end{aligned} \tag{2}$$

when an intervention with a fixed efficacy e_j is implemented to a population j with population mean FEC μ_j and aggregation parameter k_j we can write the hierarchical model for the observed FECs Y'_{ij} after an intervention as,

$$\begin{aligned} X_{ij} &\sim NB\left(\mu_j,\ k_j\right) \\ X'_{ij} &= \left(1 - e_j\right) \cdot X_{ij} \\ Z'_{ij} \Big| X'_{ij} &\sim Poiss\left(\lambda'_{ij}\right) \\ &\sim Poiss\left(f_j \cdot X'_{ij}\right) \\ Y'_{ij} &= \frac{Z'_{ij}}{f_j} \end{aligned} \tag{3}$$

By analogy with Eqn (2) the expected value and variance of Y'_{ij} are equal to,

$$\begin{aligned} E\left[Y'_{ij}\right] &= \left(1 - e_j\right) \cdot \mu_j \\ Var\left[Y'_{ij}\right] &= \left(1 - e_j\right) \cdot \mu_j \cdot \left(\frac{1}{f_j} + \left(1 - e_j\right) \cdot \left(1 + \mu_j \Big/ k_j\right)\right) \end{aligned} \tag{4}$$

The derivation of the expected valued and variance of Y_{ij} and Y'_{ij} can be found in Appendix A.

3.1.2 FECs based on pooled stool samples

A hierarchical model for FECs based on pooled stool samples can be described based on the models for FECs obtained by examining individual stool samples (see Eqns (1) and (3)). Compared to the model for individual stool samples, it is also necessary to define the number of samples pooled and the true FECs in these pools. For this model, we assume (1) that the pool size m_j (= number of individual samples pooled) is the same for all pools and (2) that the true FEC V_{lj} of a pool $l = 1, 2, \ldots, n_{j_{pool}}$ equals the mean of the true FECs of m_j individual stool samples from individuals $i = 1, 2, \ldots, n_{j_{ind}}$ from a population j ($n_{j_{ind}} = m_j \cdot n_{j_{pool}}$). The latter implies that we ignore any variation in FEC V_{lj} due to the pooling process. In practice, this means that for each

individual stool sample an equal amount of stool is pooled and that pools are well homogenized to avoid any heterogeneous distribution of eggs in the pool. The observed FEC U_{lj} prior to an intervention can be described as the hierarchical model below,

$$\begin{aligned} X_{ij} &\sim NB\left(\mu_j,\ k_j\right) \\ V_{lj} &= \frac{\sum_{i=1}^{m_j} X_{ij}}{m_j} \\ Z_{lj}\left|V_{lj}\right. &\sim Poiss\left(\lambda_{lj}\right) \\ &\sim Poiss\left(f_j \cdot V_{lj}\right) \\ U_{lj} &= \frac{Z_{lj}}{f_j} \end{aligned} \tag{5}$$

From this model and by analogy with Eqns (1) and (2), it can be shown that in a population j the expected value and the variance of observed FECs U_{lj} are equal to,

$$\begin{aligned} E\left[U_{lj}\right] &= \mu_j \\ Var\left[U_{lj}\right] &= \mu_j \cdot \left(\frac{1}{f_j} + \frac{1 + \mu_j \big/ k_j}{m_j}\right) \end{aligned} \tag{6}$$

After an intervention with a fixed efficacy e_j, we can write the hierarchical model for the observed FECs U'_{lj} in a pool of m_j stool samples from a population j as,

$$\begin{aligned} X_{ij} &\sim NB\left(\mu_j,\ k_j\right) \\ X'_{ij} &= \left(1 - e_j\right) \cdot X_{ij} \\ V'_{lj} &= \frac{\sum_{i=1}^{m_j} X'_{ij}}{m_j} \\ Z'_{lj}\left|V'_{lj}\right. &\sim Poiss\left(\lambda'_{lj}\right) \\ &\sim Poiss\left(f_j \cdot V'_{lj}\right) \\ U'_{lj} &= \frac{Z'_{lj}}{f_j} \end{aligned} \tag{7}$$

The expected value and the variance of observed FECs U'_{lj} after an intervention are equal to,

$$E\left[U'_{lj}\right] = \left(1 - e_j\right) \cdot \mu_j$$
$$Var\left[U'_{lj}\right] = \left(1 - e_j\right) \cdot \mu_j \cdot \left(\frac{1}{f_j} + \left(1 - e_j\right) \cdot \frac{1 + \mu_j \Big/ k_j}{m_j}\right) \tag{8}$$

The derivations of the expected value and the variance of U_{lj} and U'_{lj} are provided in Appendix B.

3.1.3 Distribution of $\overline{Y_{\cdot j}}$ and $\overline{U_{\cdot j}}$

Given the central limit theorem, it is expected that the mean $\overline{Y_{\cdot j}}\left(= \frac{\sum_{i=1}^{N_{j_{ind}}} Y_{ij}}{N_{j_{ind}}}\right)$ and $\overline{U_{\cdot j}}\left(= \frac{\sum_{l=1}^{N_{j_{pool}}} U_{lj}}{N_{j_{pool}}}\right)$ would approach a normal distribution when sample size is large. However, as illustrated in Figure 3, this assumption is violated when sample size is small. This figure illustrates the distribution of estimated population mean FEC for *A. lumbricoides* generated from 10,000 theoretical surveys in which 50 subjects $(= N_{j_{ind}})$ from a population j ($\mu_j = 500$ EPG and $k_j = 0.1$) are randomly screened. The mean FECs are obtained by examining both individual and pooled ($m_j = 5$, $N_{j_{pool}} = 10$) stool samples using a single Kato-Katz thick smear ($f_j = 0.0417$ g). For both examination of individual and pooled samples the distribution of $\overline{Y_{\cdot j}}$ and $\overline{U_{\cdot j}}$ is skewed, and this will become more pronounced when the sample size, μ_j and f_j decrease, and the level of aggregation increases ($\sim 1/k_j$) (data not shown). For the described hierarchical models we will assume that both $\overline{Y_{\cdot j}}$ and $\overline{U_{\cdot j}}$ follow a gamma distribution. This distribution is defined by two parameters, i.e. the shape parameter γ and the scale parameter θ. The expected value for gamma distributed variables equals $\gamma \cdot \theta$, the variance equals $\gamma \cdot \theta^2$. From these equations and given Eqns (2) and (6) we can derive γ_j and θ_j of the gamma distribution for $\overline{Y_{\cdot j}}$ Eqn (9) and $\overline{U_{\cdot j}}$ Eqn (10) for a population j given a sample size $N_{j_{ind}}$ and $N_{j_{pool}}$, respectively. As illustrated in Figure 3, these gamma distributions approach those of the above described theoretical surveys.

$$\gamma_{j_{ind}} = \frac{\mu_j \cdot N_{j_{ind}}}{1/f_j + 1 + \mu_j/k_j}$$
$$\theta_{j_{ind}} = \frac{1/f_j + 1 + \mu_j/k_j}{N_{j_{ind}}} \quad (9)$$

$$\gamma_{j_{pool}} = \frac{\mu_j \cdot N_{j_{pool}}}{1/f_j + \left(1 + \mu_j/k_j\right)/m_j}$$
$$\theta_{j_{pool}} = \frac{1/f_j + \left(1 + \mu_j/k_j\right)/m_j}{N_{j_{pool}}} \quad (10)$$

To verify whether these theoretical gamma distributions also fit empirical data, we have plotted the probability density of both $\overline{Y_{\cdot j}}$ generated by bootstrap analysis (10,000 iterations) of 108 individual data obtained by a single Kato-Katz thick smear in school ID1175 and the expected gamma distribution in Figure 5 for each of the STH species. Overall, the plots suggest that the theoretically derived gamma distributions for $\overline{Y_{\cdot j}}$ fit the empirical data. The distribution of $\overline{Y'_{\cdot j}}$ and $\overline{U'_{\cdot j}}$ will not be discussed here, but will be incorporated in Section 3.2.2, Distribution of different scenarios of $FECR_j$.

3.2 Reduction in FECs

There is an ongoing debate as to whether reduction in prevalence is an appropriate metric to monitor the impact of an intervention, as opposed to reduction in FECs (FECR syn. egg reduction rate), respectively (Humphries et al., 2011; Montresor, 2011; Montresor et al., 2011). Anderson and colleagues, highlighted that a drop in FEC may not always be reflected in a drop in prevalence, with changes in the latter hence underestimating the impact of control interventions (Anderson et al., 2012). Analogously, an intervention may fail to cure helminth infections (reduction in prevalence = 0%), but may result in an FECR of 99% which is satisfactory. Additionally, it has been shown that estimates of reduction in prevalence are highly affected by both sampling and diagnostic effort, being underestimated when the sampling and diagnostic effort are minimized. This is in sharp contrast with FECR estimates, which remain unchanged regardless of both sampling and diagnostic effort (Levecke et al., 2014a). To date, a wide range of formulae has been used to calculate FECR, each differing in terms of the statistical unit (individual vs group) and how the mean

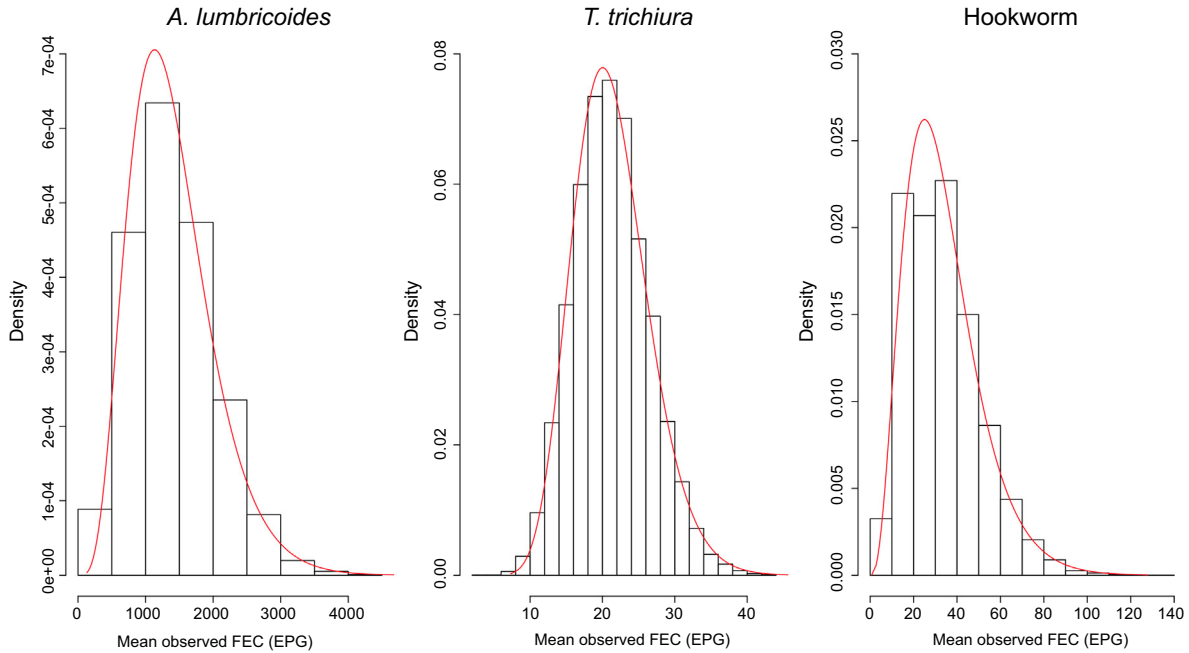

Figure 5 The histograms illustrate the probability density of mean observed faecal egg counts (FECs; expressed in eggs per gram of stool (EPG)) $(= \overline{Y_{\cdot j}})$ for *Ascaris lumbricoides*, *Trichuris trichiura* and hookworm generated by bootstrap analysis (10,000 iterations) of 108 individual data obtained by a single Kato-Katz thick smear in school ID1175. The red (light gray in print versions) lines describe the corresponding probability density of the expected gamma distribution. For these gamma distributions $\gamma_{j_{ind}}$ equalled $\frac{\mu_j \cdot N_{j_{ind}}}{1/f_j+1+\mu_j/k_j}$, $\theta_{j_{ind}}$ equalled $\frac{1/f_j+1+\mu_j/k_j}{N_{j_{ind}}}$. For each of the three helminth species, μ_j and k_j were substituted by their corresponding $\overline{Y_{\cdot j}}$ and $\overline{Y_{\cdot j}}/(Var[Y_{ij}]/\overline{Y_{\cdot j}} - 1/f_j - 1)$ (derived from Eqn (2)), respectively. For all helminth species, $N_{j_{ind}}$ was substituted by 108 and f_j by 0.0417.

FEC is calculated (arithmetic vs geometric mean). However, recent studies suggest that the group-based formula using the arithmetic mean, as described below, is a suitable metric for evaluating an intervention. Compared to the other formulae, it represents a robust indicator (vs individual-based formula) (Vercruysse et al., 2011) that provides accurate efficacy of an intervention (vs group-based formula using the geometric mean (Dobson et al., 2009; Vercruysse et al., 2011)). We will work out the applications for the assessment of FECR based on the group-based formula using the arithmetic mean.

$$FECR = 1 - \frac{arithmetic\ mean\ of\ FECs\ after\ an\ intervention}{arithmetic\ mean\ of\ FECs\ prior\ an\ intervention} \tag{11}$$

3.2.1 Expected value and variance of different FECR scenarios

When an intervention with efficacy e_j is assessed in a population j (population mean FEC of μ_j and aggregation of k_j) by means of $FECR_j$ based on the individual examination of f_j g of stool of the same individuals, we can deduce that the expected value asymptotically equals,

$$\begin{aligned} E\left[FECR_{j_{ind}}\right] &\simeq 1 - \frac{E\left[Y'_{ij}\right]}{E\left[Y_{ij}\right]} \\ &\simeq 1 - \frac{\left(1 - e_j\right) \cdot \mu_j}{\mu_j} \\ &\simeq e_j \end{aligned} \tag{12}$$

Given that two variables are dependent (examination of the same subjects both prior and after an intervention) and that both their expected value is nonzero, one can demonstrate, based on Casella and Berger (2002), that the variance of the ratio of these variables asymptotically equals,

$$\begin{aligned} Var\left[FECR_{j_{ind}}\right] \simeq \left(\frac{E\left[Y'_{ij}\right]}{E\left[Y_{ij}\right]}\right)^2 \cdot \left(\frac{Var\left[Y_{ij}\right]}{E\left[Y_{ij}\right]^2} + \frac{Var\left[Y'_{ij}\right]}{E\left[Y'_{ij}\right]^2}\right. \\ \left. - 2 \cdot corr\left(Y_{ij}, Y'_{ij}\right) \cdot \frac{\sqrt{Var\left[Y_{ij}\right] \cdot Var\left[Y'_{ij}\right]}}{E\left[Y_{ij}\right] \cdot E\left[Y'_{ij}\right]}\right) \end{aligned} \tag{13}$$

In the scenario that different subjects are examined after the intervention, the expected value remains unchanged, but the variance will change into,

$$Var\left[FECR_{j_{ind\Delta}}\right] \simeq \left(\frac{E\left[Y'_{ij}\right]}{E\left[Y_{ij}\right]}\right)^2 \cdot \left(\frac{Var\left[Y_{ij}\right]}{E\left[Y_{ij}\right]^2} + \frac{Var\left[Y'_{ij}\right]}{E\left[Y'_{ij}\right]^2}\right) \tag{14}$$

In analogy with Eqns (12)–(14), it is possible to write the expected value and the variance of $FECR_j$ based on the examination of f_j g of m_j pooled stool samples as,

$$E\left[FECR_{j_{pool}}\right] \simeq e_j \tag{15}$$

$$Var\left[FECR_{j_{pool}}\right] \simeq \left(\frac{E\left[U'_{lj}\right]}{E\left[U_{lj}\right]}\right)^2 \cdot \left(\frac{Var\left[U_{lj}\right]}{E\left[U_{lj}\right]^2} + \frac{Var\left[U'_{lj}\right]}{E\left[U'_{lj}\right]^2} - 2 \cdot corr\left(U_{lj}, U'_{lj}\right) \cdot \frac{\sqrt{Var\left[U_{lj}\right] \cdot Var\left[U'_{lj}\right]}}{E\left[U_{lj}\right] \cdot E\left[U'_{lj}\right]}\right) \tag{16}$$

$$Var\left[FECR_{j_{pool\Delta}}\right] \simeq \left(\frac{E\left[U'_{lj}\right]}{E\left[U_{lj}\right]}\right)^2 \cdot \left(\frac{Var\left[U_{lj}\right]}{E\left[U_{lj}\right]^2} + \frac{Var\left[U'_{lj}\right]}{E\left[U'_{lj}\right]^2}\right) \tag{17}$$

For each of the variances of $FECR_j$ described above it is possible to substitute the expected values and variances of Y_{ij}, Y'_{ij}, U_{lj}, U'_{lj} by the equations described in (2), (4) and (6) and (8), respectively. It was not possible to write an analytical solution for the correlation between Y_{ij} and Y'_{ij}, and between U_{lj} and U'_{lj}. Henceforth, we will estimate these correlations based on simulations.

3.2.2 Distribution of different FECR scenarios

Given the central limit theorem, it is to be expected that $FECR_j$ estimates would approach a normal distribution when sample size is large. However, as illustrated in Figure 6, this assumption is violated when sample size is small. This figure illustrates the probability density of $FECR_j$ generated from 10,000 theoretical surveys in which 50 subjects ($= N_{j_{ind}}$) from a population j ($\mu_j = 500$ EPG and $k_j = 0.1$) are included to assess the impact of three interventions e_j with three different levels of efficacy (0.50, 0.80 and 0.99) against *A. lumbricoides*. The $FECR_j$ are obtained by examining both individual and pooled stool samples ($m_j = 5, N_{j_{pool}} = 10$) using a single Kato-Katz

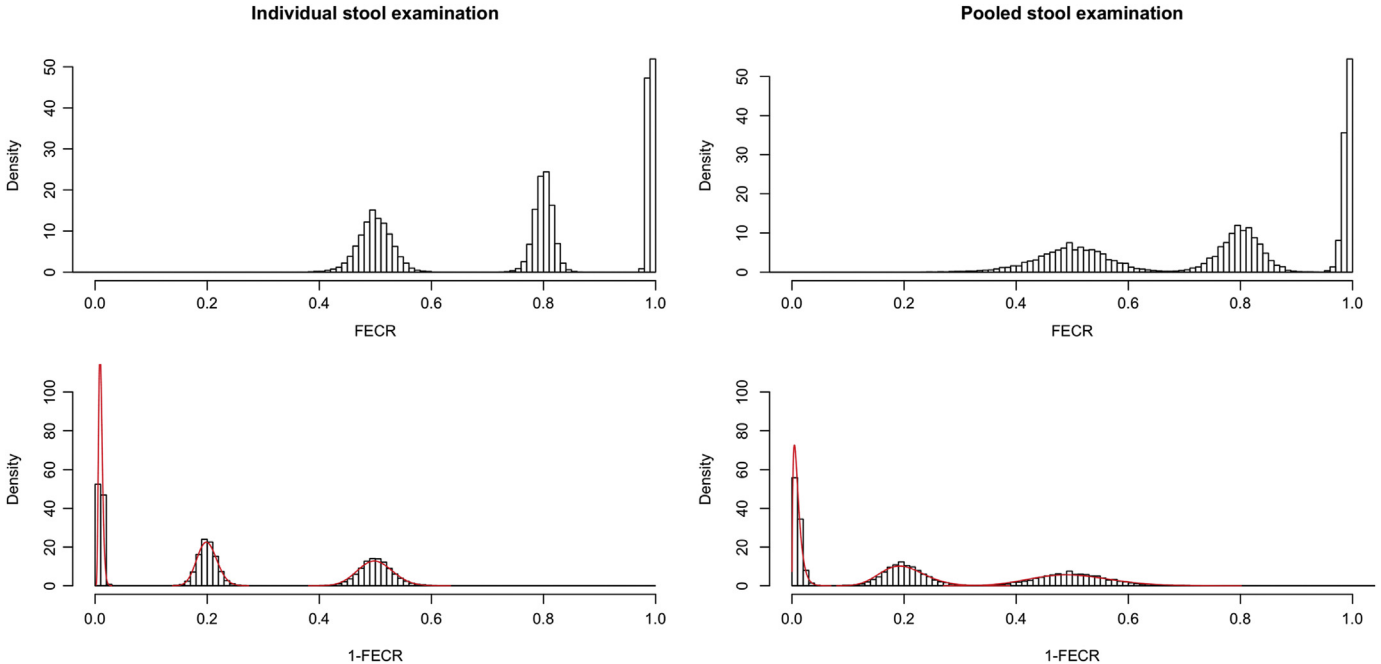

Figure 6 The top graphs illustrate the probability density of $FECR_j$ (faecal egg count reduction) estimates generated from 10,000 theoretical surveys in which 50 subjects $(= N_{j_{ind}})$ from a population j ($\mu_j = 500$ EPG and $k_j = 0.1$) are included to assess the impact of three interventions e_j with three different levels of efficacy (0.50, 0.80 and 0.99) against *Ascaris lumbricoides*. The $FECR_j$ are obtained by examining both individual (left top graph) and pooled $(m_j = 5, N_{j_{pool}} = 10)$ stool samples (right top graph) using a single Kato-Katz thick smear ($f_j = 0.0417$ g). The same individuals are screened both prior and after the intervention. The graphs at the bottom illustrate the probability density of $1 - FECR_j$ from the 10,000 theoretical surveys described above (histograms) and the probability density of the corresponding expected gamma distribution (red (light gray in print versions) line). For these gamma distributions $\gamma_{j_{ind}} = (1-e_j)^2 \cdot N_{j_{ind}} / Var[FECR_{j_{ind}}]$ and $\theta_{j_{ind}} = Var[FECR_{j_{ind}}] / (1-e_j) \cdot N_{j_{ind}}$ (left bottom graph), and $\gamma_{j_{pool}} = (1-e_j)^2 \cdot N_{j_{pool}} / Var[FECR_{j_{pool}}]$ and $\theta_{j_{pool}} = Var[FECR_{j_{pool}}] / (1-e_j) \cdot N_{j_{pool}}$.

thick smear ($f_j = 0.0417$ g). For both examination of individual and pooled samples the distribution of $FECR_j$ becomes skewed when e_j approaches 1.

$FECR_j$ can take any value from $-\infty$ (mean FEC at follow-up > mean FEC at baseline) up to 1 (mean FEC at follow-up = 0), which makes it not trivial to identify the most appropriate distribution. Instead we will focus on the function $1 - FECR_j$. This function can take any value between zero and $+\infty$, and hence we will assume that the function $1 - FECR_j$ follows a gamma distribution. The expected value of this function equals $1 - e_j$, the variance equals that of $FECR_j$ (Eqns (13), (14), (16) and (17)). Based on these equations, the general format to calculate γ_{1-FECR_j} and θ_{1-FECR_j} equals,

$$\gamma_{1-FECR_j} = \frac{\left(1 - e_j\right)^2 \cdot N_j}{Var\left[FECR_j\right]}$$
$$\theta_{1-FECR_j} = \frac{Var\left[FECR_j\right]}{\left(1 - e_j\right) \cdot N_j} \tag{18}$$

To obtain γ_{1-FECR_j} and θ_{1-FECR_j} for the different $FECR_j$ scenarios, it is necessary to substitute $Var[FECR_j]$ by Eqn (13) for $FECR_{j_{ind}}$, by Eqn (14) for $FECR_{j_{ind\Delta}}$, by Eqn (16) for $FECR_{j_{pool}}$, and by Eqn (17) for $FECR_{j_{pool\Delta}}$. As illustrated in Figure 6, these gamma distributions approach those of the above described theoretical surveys.

To verify whether these theoretical gamma distributions also fit empirical data, we have plotted the probability density of both the $1 - FECR_j$ generated by bootstrap analysis (10,000 iterations) of individual data obtained in a drug efficacy trial in Ethiopia and the expected gamma distributions in Figure 7 for each of the STH species. The drug efficacy trial was part of a multinational study designed to assess the efficacy of a single oral dose of albendazole (ABZ) against STH infections in school children in endemic countries (Vercruysse et al., 2011). For this validation all 410 ($= N_{j_{ind}}$) Ethiopian subjects screened at baseline were included. However, 154 subjects with a baseline FEC of 0 EPG were not treated nor re-examined at follow-up. To include these subjects it was assumed that the FEC at follow-up of these noninfected subjects (falsely/truly) also equalled zero after drug administration. In addition to this, seven infected subjects did not provide a stool sample at follow-up. These subjects were replaced by a random sample of subjects for which complete data were available. The values for e_j, μ_j and k_j required to estimate the parameters of the gamma distribution were estimated from the observed data. The amount of stool

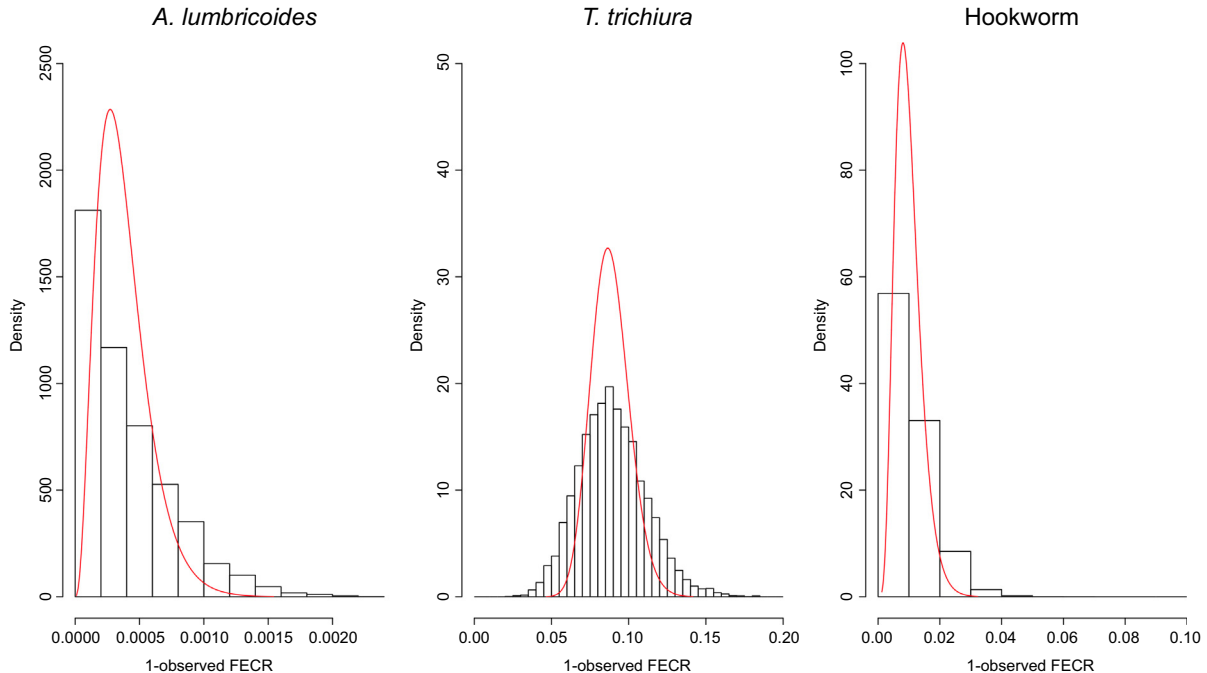

Figure 7 The probability density of $1 - FECR_j$ (faecal egg count reduction) estimates of a single oral dose of albendazole against *Ascaris lumbricoides*, *Trichuris trichiura* and hookworm infections generated by bootstrap analysis (10,000 iterations) of 410 individual data using McMaster during a drug efficacy trial in Ethiopia and the probability density of the expected gamma distribution (red (light gray in print versions) lines). In these Gamma distributions $\gamma_{j_{ind}}$ equalled $(1-e_j)^2 \cdot N_{j_{ind}} / Var[FECR_{j_{ind}}]$ and $\theta_{j_{ind}}$ equalled $Var[FECR_{j_{ind}}]/(1-e_j) \cdot N_{j_{ind}}$. For each of the three helminth species e_j, μ_j and k_j were separately substituted by $FECR_{j_{ind}}$, $\overline{Y_{\cdot j}}$ and $\overline{Y_{\cdot j}}/(Var[Y_{ij}]/\overline{Y_{\cdot j}} - 1/f_j - 1)$ (derived from Eqn (2)), respectively. For all helminth species, $N_{j_{ind}}$ was substituted by 410, and f_j by 0.02.

examined f_j was set at 0.02 g as all samples were processed with McMaster. Overall, the plots indicate that there is no perfect match between the empirical data and the theoretical gamma distributions, the theoretical distribution underestimating the variation observed in the empirical data.

4. APPLICATIONS OF THE GENERAL MATHEMATICAL FRAMEWORK

4.1 Sample size calculation for the assessment of FECs

Depending on the objectives of the survey, there are various ways to determine the samples size for the assessment of population mean FECs. We will discuss how to calculate a sample size for assessing population mean FEC with a predefined level of precision, whether population mean FEC exceeds a predefined level of infection intensity using a lot quality assurance sampling strategy (LQAS), and the absence of STH with a predefined level of probability.

4.1.1 Assessment of population mean FEC with a predefined precision

Given the variance derived from two hierarchical models (see Eqns (2) and (6)), and the distributions of $\overline{Y_{\cdot j}}$ Eqn (9) and $\overline{U_{\cdot j}}$ Eqn (10), it is possible to determine the sample size N_j to assess the population mean FEC μ_j with a precision W (= width of the $1-\alpha$ confidence interval (CI)). This can be obtained by taking the $\alpha/2th$ and $1-\alpha/2th$ percentile of the gamma distribution and determining W for a wide range of values of N_j. The required sample size is that N_j for which W does not exceed a predefined value. Figure 8 illustrates the increase in sample size as a function of W to assess the population mean of *A. lumbricoides* FEC in a population j ($\mu_j = 500$ EPG and $k_j = 0.1$). Both individual and pooled stool samples were examined using the Kato-Katz thick smear ($f_j = 0.0417$ g). In this example α was set at 0.05. When individual samples are examined, a minimum of 37 individuals ($= N_{j_{ind}}$) will need to be screened to assess the population mean FEC with a precision of at least 1,000 EPG, when pools of 5 are examined 40 individuals will need to be screened ($= N_{j_{pool}} \cdot m_j = 8 \cdot 5$).

4.1.2 Assessment of FECs using an LQAS strategy

LQAS was initially developed to assure the quality in industrial processes at minimum cost and involves three consecutive steps. First, a random sample ('lot') of items is taken, subsequently it is verified whether the lot meets a

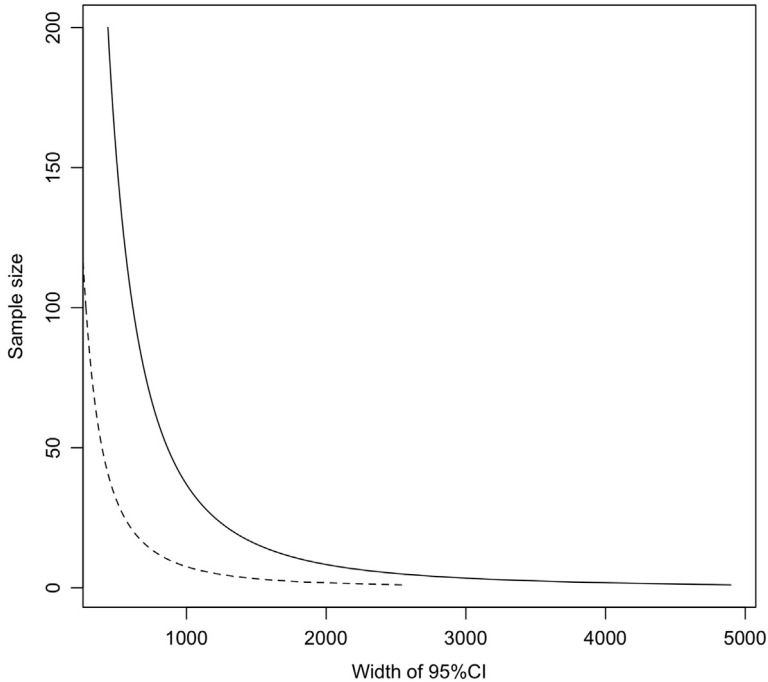

Figure 8 The number of individuals ($= N_{j_{ind}}$, straight line) and pools of 5 ($N_{j_{pool}}$, dashed line) as a function of the width of the 95% confidence interval ($\alpha = 0.05$) for the assessment of population mean *Ascaris lumbricoides* FEC in a population j ($\mu_j = 500$ EPG and $k_j = 0.1$). Both individual and pooled stool samples were examined using the Kato-Katz thick smear ($f_j = 0.0417$ g).

predefined standard of quality (e.g. number of items with production faults). Finally, further actions are taken when the quality is revealed to be unacceptable. Since its first description in the late 1920s, LQAS has also found various applications in the field of public health. For helminthiasis, it can be used to verify whether the observed number of infected subjects in a random sample exceeds a predefined decision threshold (Addiss et al., 2003; Brooker et al., 2005), followed by administration of drugs if this is the case. The mathematical underpinnings of LQAS based on helminth prevalence have been described by Olives et al. (2012). At present, LQAS has not yet been developed for FECs.

For the application of LQAS based on FECs, we want to identify those populations j for which μ_j equals or exceeds a decision threshold t, and hence those populations j that are at the highest need of drug

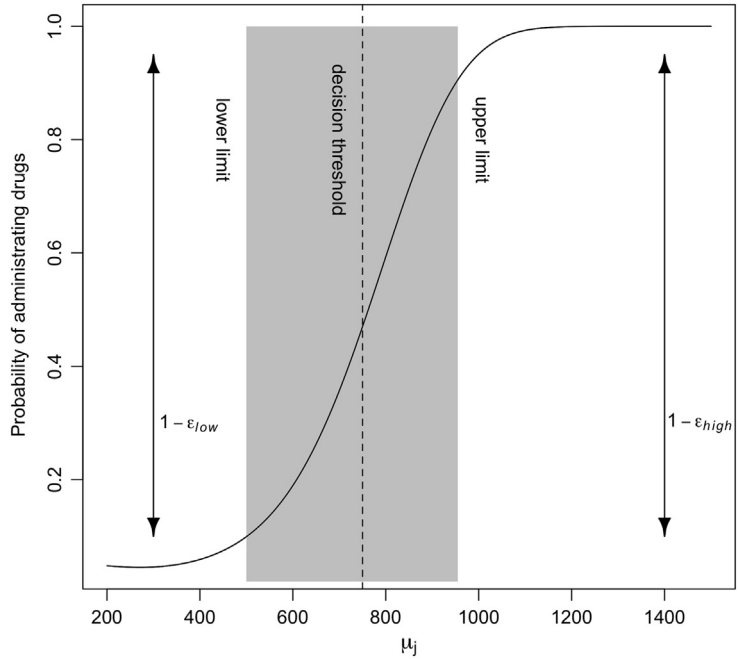

Figure 9 The probability of administering drugs to a population *j* as a function of population mean faecal egg count μ_j (expressed in eggs per g of stool). The dashed line indicates the decision threshold *t*. The grey zone between the lower and the upper limit represents the range of μ_j values for which the probability of correctly withdrawing drugs $(= 1 - \varepsilon_{low_j})$ or correctly administering drugs $(= 1 - \varepsilon_{high_j})$ is less than 0.90 $(\varepsilon_{low_j} = \varepsilon_{high_j} = 0.1)$.

administration. However, we want to reduce the probability (ε_{low_j}) that populations are unnecessarily receiving drugs and that populations in need of drugs are not covered (ε_{high_j}). Therefore, we want to be confident that infected populations in no need of drugs with $\mu_j \leq ll$ (lower limit) will not be covered with a probability of at least $1 - \varepsilon_{low_j}$ and that heavily infected populations with $\mu_j \geq ul$ (upper limit) will receive drugs with a probability of at least $1 - \varepsilon_{high_j}$. We illustrated the concept of LQAS in Figure 9. In this figure, the sample size N_j allowed to correctly withhold and provide drugs to a population j with a probability of at least 0.90 $(\varepsilon_{low_j} = \varepsilon_{high_j} = 0.1)$ as long as $\mu_j \leq 500$ EPG, and $\mu_j \geq 955$ EPG, respectively.

Given that $\overline{Y_{\cdot j}}$ and $\overline{U_{\cdot j}}$ are gamma distributed, we can write these probabilities for the examination of individual stool samples as,

$$\begin{aligned}
\varepsilon_{low_{j_{ind}}} &\leq P\left(\overline{Y_{\cdot j}} \geq t \middle| \mu_j = ll\right) = 1 - P\left(\overline{Y_{\cdot j}} < t \middle| \mu_j = ll\right) \\
&\leq 1 - F\left(t; \gamma_{j_{ind}}, \theta_{j_{ind}}, N_{j_{ind}} \middle| \mu_j = ll\right) \\
\varepsilon_{high_{j_{ind}}} &\leq P\left(\overline{Y_{\cdot j}} < t \middle| \mu_j = ul\right) \\
&\leq F\left(t; \gamma_{j_{ind}}, \theta_{j_{ind}}, N_{j_{ind}} \middle| \mu_j = ul\right)
\end{aligned} \tag{19}$$

and for the examination of pooled stool samples as

$$\begin{aligned}
\varepsilon_{low_{j_{pool}}} &\leq P\left(\overline{U_{\cdot j}} \geq t \middle| \mu_j = ll\right) = 1 - P\left(\overline{U_{\cdot j}} < t \middle| \mu_j = ll\right) \\
&\leq 1 - F\left(t; \gamma_{j_{pool}}, \theta_{j_{pool}}, N_{j_{pool}} \middle| \mu_j = ll\right) \\
\varepsilon_{high_{j_{pool}}} &\leq P\left(\overline{U_{\cdot j}} < t \middle| \mu_j = ul\right) \\
&\leq F\left(t; \gamma_{j_{pool}}, \theta_{j_{pool}}, N_{j_{pool}} \middle| \mu_j = ul\right)
\end{aligned} \tag{20}$$

F in these equations represents the lower tail of the cumulative distribution function of the gamma distribution. For the parameterization of $\gamma_{j_{ind}}$, $\theta_{j_{ind}}$, $\gamma_{j_{pool}}$ and $\theta_{j_{pool}}$, we refer the reader to Eqns (9) and (10). Subsequently, the required sample size N_j can be determined for which ε_{low_j} at $\mu_j = ll$ and ε_{high_j} at $\mu_j = ul$ do not exceed a predefined value. In the following hypothetical example, we will determine the sample size N_j to correctly assign drugs against *A. lumbricoides* infections in a population j. The threshold t to administer drugs is set at 750 EPG. We would like to have 90% confidence that populations with $\mu_j = 500$ EPG are withhold from treatment, and that populations with $\mu_j = 1000$ EPG receive treatment, implying that both ε_{low_j} and ε_{high_j} equal 0.1. Given that μ_j and k_j are correlated with each other (see Figure 2), we will assume that k_j equals 0.1 and 0.6 for a population j for which μ_j equals 500 EPG and 1,000 EPG, respectively. We will screen both individual and pooled samples ($m_j = 5$) using a single Kato-Katz thick smear ($f_j = 0.0417$ g). Figure 10 illustrates the sample size as a function of ε_{low_j} and ε_{high_j}. When samples are examined individually at least 72 individuals are required (N_{ind} for $\varepsilon_{low_{j_{ind}}} \leq 0.1$ equals 72, $N_{j_{ind}}$ for $\varepsilon_{high_{j_{ind}}} \leq 0.1$ equals 41). When samples are pooled into pools of 5, 75 individuals are required ($N_{j_{pool}}$ for $\varepsilon_{low_{j_{pool}}} \leq 0.1$ equals 15, $N_{j_{pool}}$ for $\varepsilon_{high_{j_{pool}}} \leq 0.1$ equals 9).

4.1.3 Assessment of absence of STHs

When assessing the absence of STH infections, we would like to minimize the probability of falsely declaring the absence of eggs based on a subset of

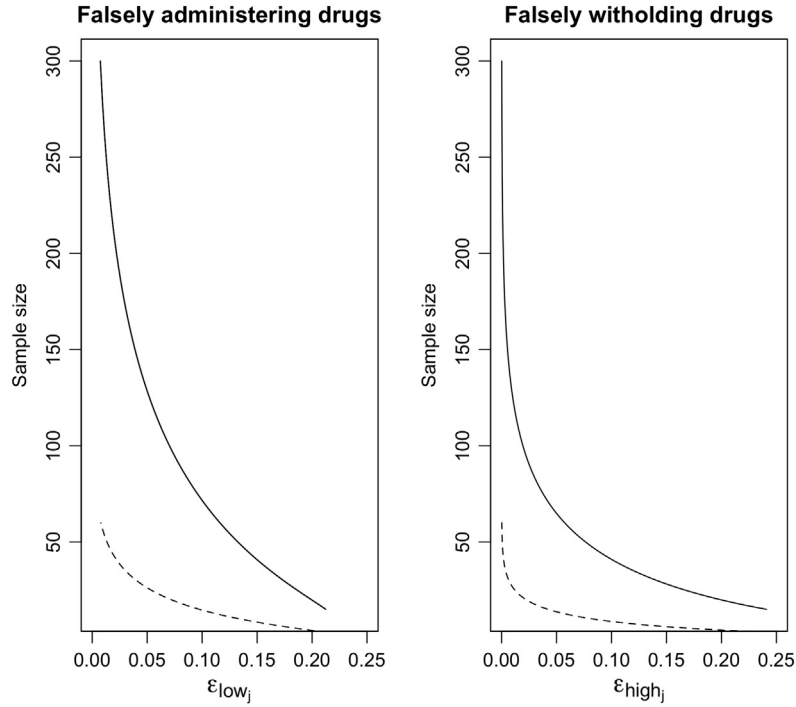

Figure 10 The number of individuals $(= N_{j_{ind}}$, straight line) and pools of 5 ($N_{j_{pool}}$, dashed line) as a function of the probability of falsely distributing drugs ε_{low_j} (left graph) in a population j ($\mu_j = 500$ EPG and $k_j = 0.1$), and falsely withholding drugs ε_{high_j} in a population j ($\mu_j = 1000$ EPG and $k_j = 0.6$) (right graph).

$N_{j_{ind}}$ individuals $(\overline{Y_{\cdot j}} = 0)$ or $N_{j_{pool}}$ pools $(\overline{U}_{\cdot j} = 0)$, while population mean FEC μ_j of a population j exceeds zero. Given that $\overline{Y_{\cdot j}}$ and $\overline{U_{\cdot j}}$ are gamma distributed, we can deduce this probability ε_{zero_j} for examination of individual stool samples as described in Eqns (21) and (22). The terms $1/(f \cdot N_{j_{ind}})$ and $1/(f \cdot N_{j_{pool}})$ in these equation correspond to the least nonzero mean FEC possible. For example, if 25 subjects $(= N_{j_{ind}})$ are screened individually for the presence of eggs using a single Kato-Katz thick smear ($f_j = 0.0417$ g), then the least nonzero possible number of eggs detected is one single egg in one subject, or a mean FEC of 1 egg/$(0.0417 \text{ g} \cdot 25) = 0.96$ EPG.

$$\begin{aligned} \varepsilon_{zero_{j_{ind}}} &= P\left(\overline{Y_{\cdot j}} = 0 \middle| \mu_j > 0\right) = P\left(\overline{Y_{\cdot j}} < 1 \middle/ \left(f_j \cdot N_{j_{ind}}\right) \middle| \mu_j > 0\right) \\ &= F\left(1 \middle/ \left(f_j \cdot N_{j_{ind}}\right); \gamma_{j_{ind}}, \theta_{j_{ind}}, N_{j_{ind}} \middle| \mu_j > 0\right) \end{aligned} \tag{21}$$

$$\begin{aligned}\varepsilon_{zero_{j_{pool}}} &= P\left(\overline{U_{\cdot j}} = 0 \middle| \mu_j > 0\right) = P\left(\overline{U_{\cdot j}} < 1 \middle/ \left(f_j \cdot N_{j_{pool}}\right) \middle| \mu_j > 0\right) \\ &= F\left(1 \middle/ \left(f_j \cdot N_{j_{pool}}\right); \gamma_{j_{pool}}, \theta_{j_{pool}}, N_{j_{pool}} \middle| \mu_j > 0\right)\end{aligned} \tag{22}$$

For the parameterization of $\gamma_{j_{ind}}$, $\theta_{j_{ind}}$, $\gamma_{j_{pool}}$, $\theta_{j_{pool}}$ we refer the reader to Eqns (9) and (10). Figure 11 illustrates the probability ε_{zero_j} to falsely declare a population *j* free of *A. lumbricoides* infections as a function of sample size N_j. In this theoretical example, the population mean FEC μ_j and aggregation parameter k_j equalled 100 EPG and 0.001, respectively. Both individual and pooled ($m_j = 5$) samples were examined using a Kato-Katz thick smear ($f_j = 0.0417$ g). If we would like to minimize ε_{zero_j} to 0.1, 290 individuals need to be screened when samples are examined individually. When pools of 5 are examined, 360 individuals need to be screened ($= N_{j_{pool}} \cdot m_j = 72 \cdot 5$).

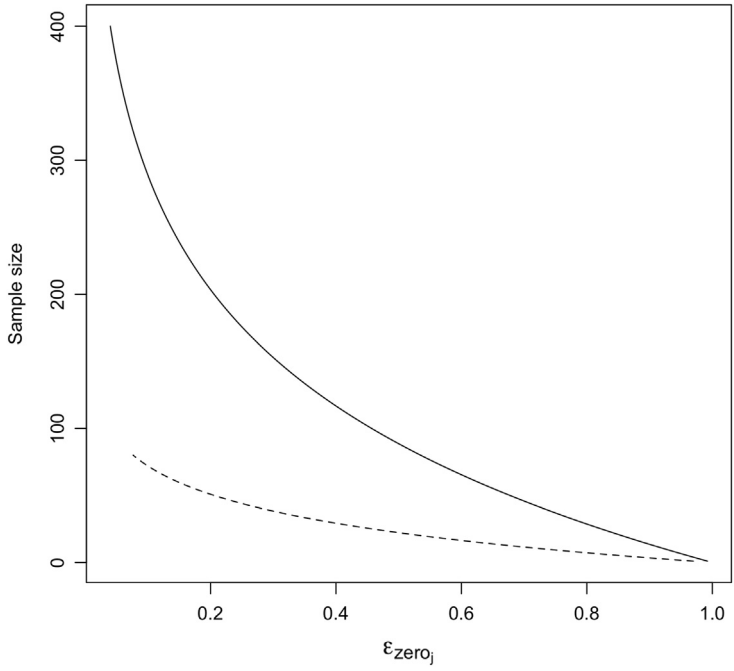

Figure 11 The number of individuals ($= N_{j_{ind}}$, straight line) and pools of 5 ($N_{j_{pool}}$, dashed line) as a function of the probability ε_{zero_j} of falsely declaring that a population *j* is free of *Ascaris lumbricoides* infections, while the population mean FEC μ_j and aggregation parameter k_j equals 100 eggs per gram of stool and 0.001, respectively. Both individual samples and pooled samples were examined with a Kato-Katz thick smear ($f_j = 0.0417$ g).

4.2 Sample size calculation for the assessment of FECR

4.2.1 Assessment of FECR with a predefined precision

We will determine the sample size that allows to assess e_j with a precision W (= width of the $1 - \alpha$ CI). This can be determined by taking the $\alpha/2$th and $1-\alpha/2$th percentile of the gamma distribution and determine W for a wide range of values of N_j. The required sample size is that N_j for which W does not exceed a predefined value. Figure 12 illustrates the increase in sample size as a function of W to assess the efficacy of an intervention e_j (= 98%) in a population j ($\mu_j = 500$ EPG and $k_j = 0.1$). We will re-examine the same subjects after intervention and both individual and pooled stool samples will be examined using the Kato-Katz thick smear ($f_j = 0.0417$ g). In this example, α was set at 0.05. When individual samples are examined a minimum of 14 individuals $(= N_{j_{ind}})$ will need to be screened to assess

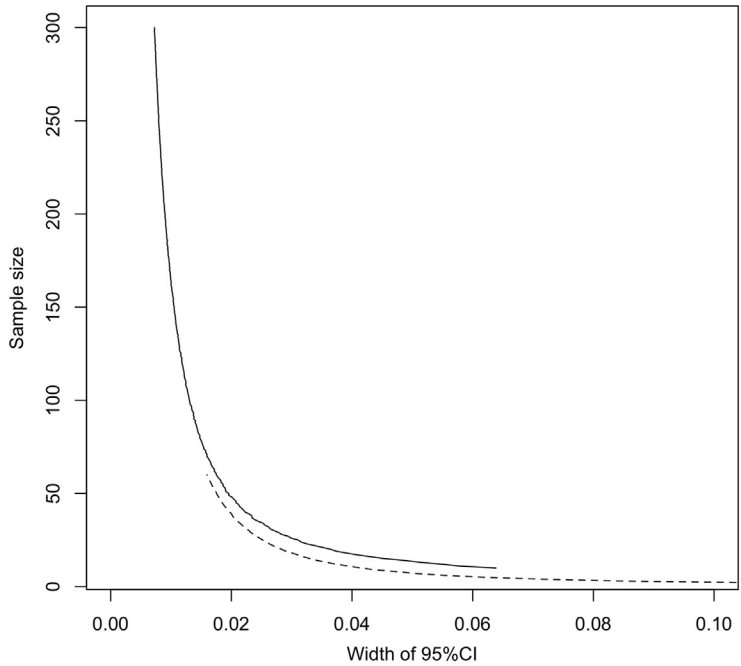

Figure 12 The number of individuals $(= N_{j_{ind}}$, straight line) and pools of 5 ($N_{j_{ind}}$, dashed line) as a function of the width of the 95% confidence interval ($\alpha = 0.05$) for the assessment of efficacy of an intervention e_j (= 0.98) based on faecal egg count reduction (FECR) in population j ($\mu_j = 500$ EPG and $k_j = 0.1$). Both individual samples ($FECR_{j_{ind}}$) and pooled samples ($FECR_{j_{pool}}$) were examined with a Kato-Katz thick smear ($f_j = 0.0417$ g). The same individuals were screened both prior to and after the intervention.

FECR with a precision of at least 0.05, when pools of 5 are examined 40 individuals will need to be screened ($= N_{j_{pool}} \cdot m_j = 8 \cdot 5$). In addition to the number of subjects that is required to guarantee a reliable assessment of e_j of an intervention, it is necessary also to ensure that e_j can be estimated at all. In the case of $\overline{Y_{\cdot j}} = 0$ or $\overline{U_{\cdot j}} = 0$, this will not be not possible. We will therefore calculate the sample size required that permits to assess e_j in a population j ($\overline{Y_{\cdot j}} > 0$ or $\overline{U_{\cdot j}} > 0$) with a probability of at least $1 - \varepsilon_{zero_j} \cdot \varepsilon_{zero_j}$ will be determined as described in Section 4.1.3, Assessment of absence of STH. For these sample sizes obtained to assess $FECR_j$ with a $W = 0.05$, the probability of ε_{zero_j} for examination of individual samples is less than 0.0005, for examination of pooled samples this was less than 10^{-7}. To obtain a comparable level of ε_{zero_j}, 28 individual subjects will need to be screened.

4.2.2 Assessment of FECR using an LQAS strategy

An LQAS strategy to assess FECR is most appropriate to verify whether the efficacy of drugs is still satisfactory. Based on two multicentre studies assessing the efficacy of ABZ (Vercruysse et al., 2011) and mebendazole (MBZ) (Levecke et al., 2014b), the WHO has defined the expected minimum for these drugs in all future drug efficacy studies. Drug efficacies below these levels should be viewed as indicative of potential drug resistance (WHO, 2013). For both drugs, the decision thresholds t are set at 0.95 for *A. lumbricoides* and 0.50 for *T. trichiura*. For hookworms, the threshold is set at 0.90 when ABZ is administered and 0.70 when MBZ is administered. For analogy with assigning a treatment to a population j when μ_j exceeds a threshold t, it is possible to write the probability $\varepsilon_{satisfact_j}$ of falsely concluding that a drug has satisfactory efficacy and the probability $\varepsilon_{reduced_j}$ of falsely concluding that the drug has reduced efficacy. For the examination of individual stool samples of the same individuals both prior to and after the intervention, these probabilities can be written as,

$$\begin{aligned}
\varepsilon_{satisfact_{j_{ind}}} &\leq P\left(FECR_{j_{ind}} \geq t \middle| e_j = ll\right) \\
&\leq P\left(1 - FECR_{j_{ind}} \leq 1 - t \middle| e_j = ll\right) \\
&\leq F\left(1 - t; \gamma_{j_{ind}}, \theta_{j_{ind}}, N_{j_{ind}} \middle| e_j = ll\right)
\end{aligned} \tag{23}$$

$$\begin{aligned}
\varepsilon_{reduced_{j_{ind}}} &\leq P\left(FECR_{j_{ind}} < t \middle| e_j = ul\right) = P\left(1 - FECR_{j_{ind}} > 1 - t \middle| e_j = ul\right) \\
&\leq 1 - P\left(1 - FECR_{j_{ind}} \leq 1 - t \middle| e_j = ul\right) \\
&\leq 1 - F\left(1 - t; \gamma_{j_{ind}}, \theta_{j_{ind}}, N_{j_{ind}} \middle| e_j = ul\right)
\end{aligned}$$

for the examination of individual stool samples of different subjects prior to and after the intervention as,

$$\begin{aligned}
\varepsilon_{satisfact_{j_{ind\Delta}}} &\le P\left(FECR_{j_{ind\Delta}} \ge t \middle| e_j = ll\right) \\
&\le P\left(1 - FECR_{j_{ind\Delta}} \le 1 - t \middle| e_j = ll\right) \\
&\le F\left(1 - t; \gamma_{j_{ind\Delta}}, \theta_{j_{ind\Delta}}, N_{j_{ind\Delta}} \middle| e_j = ll\right) \\
\varepsilon_{reduced_{j_{ind\Delta}}} &\le P\left(FECR_{j_{ind\Delta}} < t \middle| e_j = ul\right) = P\left(1 - FECR_{j_{ind\Delta}} > 1 - t \middle| e_j = ul\right) \\
&\le 1 - P\left(1 - FECR_{j_{ind\Delta}} \le 1 - t \middle| e_j = ul\right) \\
&\le 1 - F\left(1 - t; \gamma_{j_{ind\Delta}}, \theta_{j_{ind\Delta}}, N_{j_{ind\Delta}} \middle| e_j = ul\right)
\end{aligned} \tag{24}$$

for the examination of pooled stool samples of the same subjects prior to and after the intervention as,

$$\begin{aligned}
\varepsilon_{satisfact_{j_{pool}}} &\le P\left(FECR_{j_{pool}} \ge t \middle| e_j = ll\right) \\
&\le P\left(1 - FECR_{j_{pool}} \le 1 - t \middle| e_j = ll\right) \\
&\le F\left(1 - t; \gamma_{j_{pool}}, \theta_{j_{pool}}, N_{j_{pool}} \middle| e_j = ll\right) \\
\varepsilon_{reduced_{j_{pool}}} &\le P\left(FECR_{j_{pool}} < t \middle| e_j = ul\right) = P\left(1 - FECR_{j_{pool}} > 1 - t \middle| e_j = ul\right) \\
&\le 1 - P\left(1 - FECR_{j_{pool}} \le 1 - t \middle| e_j = ul\right) \\
&1 - F\left(1 - t; \gamma_{j_{pool}}, \theta_{j_{pool}}, N_{j_{pool}} \middle| e_j = ul\right)
\end{aligned} \tag{25}$$

and for the examination of pooled stool samples of different subjects prior to and after the intervention as,

$$\begin{aligned}
\varepsilon_{satisfact_{j_{pool\Delta}}} &\le P\left(FECR_{j_{pool\Delta}} \ge t | e = ll\right) \\
&\le P\left(1 - FECR_{j_{pool\Delta}} \le 1 - t \middle| e_j = ll\right) \\
&\le F\left(1 - t; \gamma_{j_{pool\Delta}}, \theta_{j_{pool\Delta}}, N_{j_{pool\Delta}} \middle| e_j = ll\right) \\
\varepsilon_{reduced_{j_{pool\Delta}}} &\le P\left(FECR_{j_{pool\Delta}} < 1 - t \middle| e_j = ul\right) = P\left(1 - FECR_{j_{pool\Delta}} > 1 - t \middle| e_j = ul\right) \\
&\le 1 - P\left(1 - FECR_{j_{pool\Delta}} \le 1 - t \middle| e_j = ul\right) \\
&\le 1 - F\left(1 - t; \gamma_{j_{pool\Delta}}, \theta_{j_{pool\Delta}}, N_{j_{pool\Delta}} \middle| e_j = ul\right)
\end{aligned} \tag{26}$$

Subsequently, the required sample size N_j can be determined for which $\varepsilon_{satisfact_j}$ at $e_j = ll$ and $\varepsilon_{reduced_j}$ at $e_j = ul$ does not exceed a predefined value. In the following hypothetical example we will determine the sample size N_j required to correctly classify the efficacy of ABZ against *A. lumbricoides* as unsatisfactory when e_j is up to 0.90 ($= ll$), and to correctly classify the efficacy as satisfactory if e is at least 0.98 (ul) with a probability of 0.90, implying that both $\varepsilon_{satisfact_j}$ and $\varepsilon_{reduced_j}$ equal 0.1. We will conduct the survey in a population j ($\mu_j = 500$ EPG and $k_j = 0.1$) and screen both individual and pooled samples ($m_j = 5$) using a single Kato-Katz thick smear ($f_j = 0.0417$ g). The same individuals are screened both prior to and after the intervention. Figure 13 illustrates the sample size as a function of $\varepsilon_{satisfact_j}$ and $\varepsilon_{reduced_j}$. When samples are examined individually at least 19 individuals are required ($N_{j_{ind}}$ for $\varepsilon_{satisfact_{j_{ind}}} \leq 0.1$ equals 19, $N_{j_{ind}}$ for $\varepsilon_{reduced_{j_{ind}}} \leq 0.1$ equals 18). When samples are pooled into pools of 5, 60 individuals are required ($N_{j_{pool}}$ for $\varepsilon_{satisfact_{j_{pool}}} \leq 0.1$ equals 11, $N_{j_{pool}}$ for $\varepsilon_{reduced_{j_{pool}}} \leq 0.1$ equals 12). For these sample sizes ε_{zero_j} for examination of individual samples is less than $2.5 \cdot 10^{-5}$, for examination of pooled samples this was less than $4.5 \cdot 10^{-13}$. To obtain a comparable level of ε_{zero_j}, 46 individual subjects will need to be screened.

4.3 Estimation of true prevalence in absence of a gold standard

Accurate assessment of true prevalence (*TP*) depends largely on the ability of the diagnostic technique to correctly identify the presence (sensitivity) and absence (specificity) of eggs in stool. The eggs excreted by STHs have a unique and distinct morphology, enabling a specific diagnosis (no false positives). However, diagnostic methods often lack sensitivity (false negatives), consequently the apparent prevalence (*AP*) of subjects excreting eggs underestimates in most cases *TP*. There is a broad literature available on how to estimate the *TP* in absence of a gold standard, of which the majority focuses on latent class models. In these models unknown (latent) categorical (class) variables are estimated based either on observed data (maximum likelihood approach) or on a combination of observed data and *a priori* scientific knowledge (Bayesian approach). It is not the scope of the current study to discuss the different estimation procedures in detail, for this we refer the reader to Hui and Zou (1998), Enøe et al. (2000), Diggle (2011) and Speybroeck et al. (2013). Instead, we would like to indicate that both approaches are fed with binary inputs only (positive or

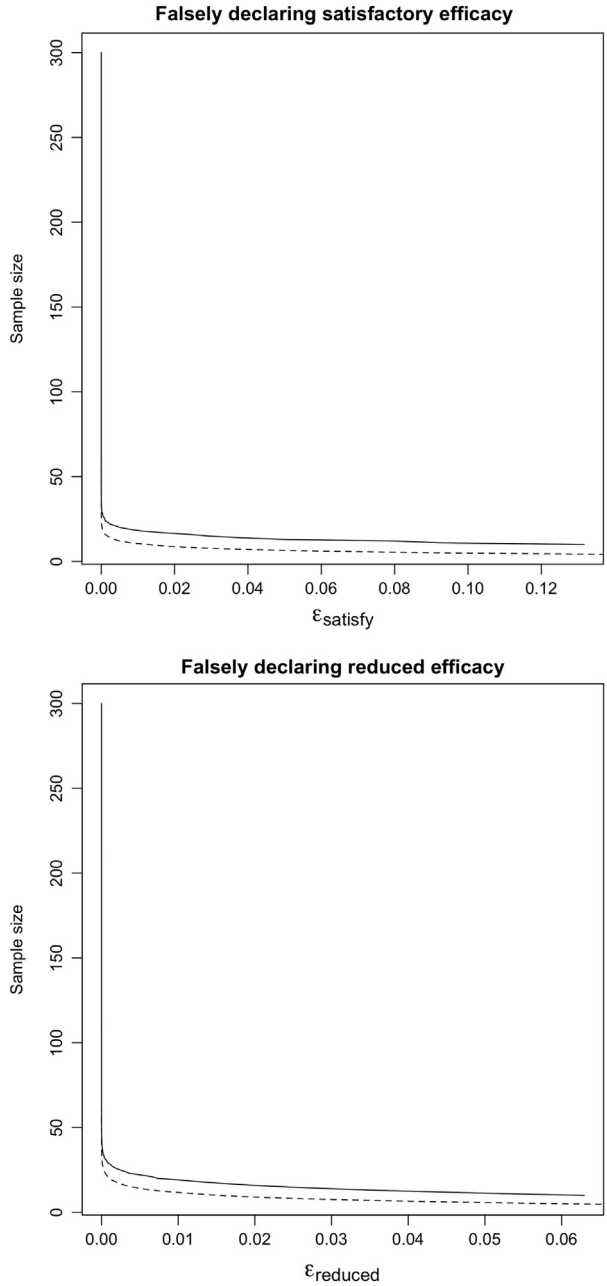

Figure 13 The number of individuals $(= N_{j_{ind}}, \text{straight line})$ and pools of 5 $(N_{j_{pool}}, \text{dashed line})$ as a function of the probability of falsely classifying drugs $\varepsilon_{satisfact_j}$ as satisfactory when $e_j = 0.90$ (top graph) and reduced $\varepsilon_{reduced_j}$ when $e_j = 0.98$ (bottom graph) with a probability not higher than 0.1 in a population j ($\mu_j = 500$ EPG and $k_j = 0.1$). Both individual samples and pooled samples were examined with a Kato-Katz thick smear ($f_j = 0.0417$ g). The same individuals are screened both prior and after the intervention.

negative test result), which are typically summarized in cross tables when at least two diagnostic methods are applied. In doing so for STHs, they totally ignore the underlying FECs which, as illustrated by Figure 4, largely explain the variation in sensitivity of diagnostic techniques between STH species, individuals and populations (see also Levecke et al., 2011b). In addition, from a programmatic point of view it is also important to periodically re-evaluate the prevalence of both any STH infection and the prevalence of low, moderate and high intensity infections after consecutive rounds of MDA. Prevalence of any STHs is currently the recommended parameter to scale down the frequency of MDA (WHO, 2011). Morbidity caused by STHs is generally more pronounced for moderate and high intensity infections, and hence a shift towards infections of low intensity is an additional parameter to evaluate the impact of MDA on public health. We will discuss how to estimate the *TP* of individual STH species, any STH and the three levels of infection intensity, taking into account the variation in sensitivity between individuals due to varying FECs.

4.3.1 Estimation of true prevalence of individual STH species

In our mathematical framework, the true prevalence TP_j of a particular STH species in a population j with mean population FEC μ_j and aggregation parameter k_j equals $1-P(X_{ij}=0)=1-\frac{\Gamma(x+k_j)}{\Gamma(k_j)\cdot x!}\cdot\left(\frac{k_j}{k_j+\mu_j}\right)^{k_j}\cdot\left(1-\frac{k_j}{k_j+\mu_j}\right)^{x}=1-\left(\frac{k_j}{k_j+\mu_j}\right)^{k_j}$.

When individual stool samples are examined, we can substitute μ_j by $\overline{Y_{\cdot j}}$; subsequently k_j can be substituted by $\overline{Y_{\cdot j}}/(Var[Y_{ij}]/\overline{Y_{\cdot j}}-1/f_j-1)$ (derived from Eqn (2)) resulting in the following equation to estimate TP_j,

$$E\left[TP_{j_{ind}}\right]=1-\left(\frac{\overline{Y_{\cdot j}}\Big/\left(Var\left[Y_{ij}\right]\Big/\overline{Y_{\cdot j}}-1\Big/f_j-1\right)}{\overline{Y_{\cdot j}}\Big/\left(Var\left[Y_{ij}\right]\Big/\overline{Y_{\cdot j}}-1\Big/f_j-1\right)+\overline{Y_{\cdot j}}}\right)^{\overline{Y_{\cdot j}}\Big/\left(Var\left[Y_{ij}\right]\Big/\overline{Y_{\cdot j}}-1\Big/f_j-1\right)} \tag{27}$$

By analogy with this equation, the expected $TP_{j_{pool}}$ based on observed FECs U_{lj} obtained by screening f_j g of pooled stool of m_j individual samples can be estimated as,

$$E\left[TP_{j_{pool}}\right] =$$

$$1 - \left(\frac{\overline{U_{\cdot j}} \Big/ \left(m_j \cdot \left(Var\left[U_{lj}\right] \Big/ \overline{U_{\cdot j}} - 1 \Big/ f_j\right) - 1\right)}{\overline{U_{\cdot j}} \Big/ \left(m_j \cdot \left(Var\left[U_{lj}\right] \Big/ \overline{U_{\cdot j}} - 1 \Big/ f_j\right) - 1\right) + \overline{U_{\cdot j}}}\right)^{\overline{U_{\cdot j}} \Big/ \left(m_j \cdot \left(Var\left[U_{lj}\right] \Big/ \overline{U_{\cdot j}} - 1 \Big/ f_j\right) - 1\right)} \tag{28}$$

In the following theoretical example, we sampled 250 subjects from an infinite population j to assess the prevalence of *A. lumbricoides*. In this population, μ_j and k_j for *A. lumbricoides* equal 500 EPG and 0.1, respectively. Both individual and pooled samples ($m_j = 10$) were screened using a single Kato-Katz thick smear ($f_j = 0.0417$ g). The histogram of the true FECs (X_{ij}), the observed individual (Y_{ij}) and pooled FECs (U_{lj}), and their respective *AP*, mean and variance are provided in Figure 14. The TP_j in this population j equals 57.3% ($= 1 - (0.1/(0.1 + 500))^{0.1}$). By substituting expected values and variances in Eqns (27) and (28) by the observed means and variances, we obtain a $TP_{j_{ind}}$ (and corresponding 95%CI) of 64.1% (33.2; 94.0) and a $TP_{j_{pool}}$ 63.2% (31.1; 98.7) based on examination of individual and pooled stool samples, respectively. The 95%CI were obtained by a bootstrap analysis which took into account the correlation between mean and variance separately for Y_{ij} and U_{lj}.

4.3.2 Estimation of true prevalence of any STH

We can determine the *TP* of any STH by combining the above described Eqns (27) and (28) with those described by de Silva and Hall (2010) that allowed to estimate the prevalence of any STH based on the prevalence estimates of the individual STH species. We adapted their formulae below. In this formula $TP_{j_{sth}}$ represents the *TP* of any STH in a population j, whereas TP_{j_a}, TP_{j_t} and TP_{j_h} represent the *TP* of *A. lumbricoides*, *T. trichiura* and hookworm, respectively. The constant 1.06 is a correction factor suggested by de Silva and Hall after validation of their formulae with field data.

$$E\left[TP_{j_{sth}}\right] = \Big(E\left[TP_{j_a}\right] + E\left[TP_{j_t}\right] + E\left[TP_{j_h}\right] - \Big(E\left[TP_{j_a}\right] \cdot E\left[TP_{j_t}\right] + E\left[TP_{j_a}\right] \cdot E\left[TP_{j_h}\right] + E\left[TP_{j_t}\right] \cdot E\left[TP_{j_h}\right] + E\left[TP_{j_a}\right] \cdot E\left[TP_{j_t}\right] \cdot E\left[TP_{j_h}\right]\Big)\Big) \Big/ 1.06 \tag{29}$$

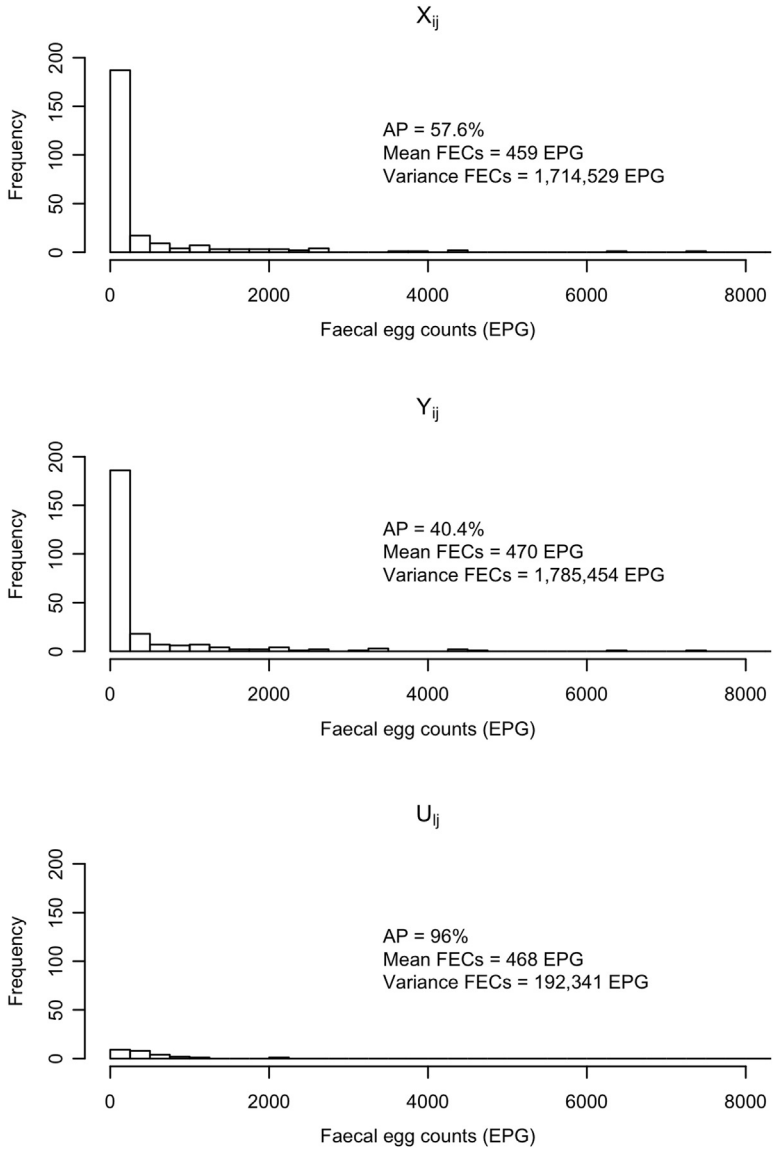

Figure 14 The histograms of the true *Ascaris lumbricoides* faecal egg counts (FECs; expressed in eggs per g of stool (EPG)) (X_{ij}), the observed individual (Y_{ij}) and pooled FECs (U_{lj}) from a random sample of 250 subjects from an infinitive population j ($\mu_j = 500$ EPG and $k_j = 0.1$). Both individual and pooled samples ($m_j = 10$) were screened using a single Kato-Katz thick smear ($f_j = 0.0417$ g). The mean and variance of the different FECs, and their corresponding apparent prevalence (*AP*) are provided.

If only two STH species are present in a population j, for example *A. lumbricoides* and *T. trichiura*, we estimate the TP of any of those two STH infections $TP_{j_{at}}$ as,

$$E\left[TP_{j_{at}}\right] = \left(E\left[TP_{j_a}\right] + E\left[TP_{j_t}\right] - E\left[TP_{j_a}\right] \cdot E\left[TP_{j_t}\right]\right)/1.06 \tag{30}$$

In analogy, the TP of any other combination of two STH species infections can be written as,

$$E\left[TP_{j_{ah}}\right] = \left(E\left[TP_{j_a}\right] + E\left[TP_{j_h}\right] - E\left[TP_{j_a}\right] \cdot E\left[TP_{j_h}\right]\right)/1.06 \tag{31}$$

for estimating the TP of any *A. lumbricoides* and hookworm infection, and as,

$$E\left[TP_{j_{th}}\right] = \left(E\left[TP_{j_t}\right] + E\left[TP_{j_h}\right] - E\left[TP_{j_t}\right] \cdot E\left[TP_{j_h}\right]\right)/1.06 \tag{32}$$

for estimating the TP of any *T. trichiura* and hookworm infection.

4.3.3 Estimation of true prevalence of low, moderate and high intensity infections

The intensity of STH infections can be classified into low, moderate and high based on the individual FECs (in EPG). The FEC thresholds for each of the STH species are proposed by the WHO (1998) and are summarized in Table 4.

Because these thresholds are STH species-specific (due to the differences in fecundity and morbidity between these species) we only will work out the framework for *A. lumbricoides*. Subsequently, a framework can be developed for the two remaining STHs by replacing the *A. lumbricoides* thresholds by those for *T. trichiura* and hookworm, respectively.

For *A. lumbricoides*, the TP of low intensity infections in a population j with mean population FEC μ_j and aggregation parameter k_j can be written as,

$$\begin{aligned} E\left[TP_{j_{low_a}}\right] &= P\left(1 \geq X_{ij} \leq 4999\right) \\ &= P\left(X_{ij} \leq 4999\right) - P\left(X_{ij} = 0\right) \\ &= F\left(4999; \mu_j, k_j\right) - F\left(0; \mu_j, k_j\right) \end{aligned} \tag{33}$$

Table 4 The faecal egg count thresholds (expressed in number of eggs in 1 g of stool) proposed by WHO (1998) to classify the intensity of *Ascaris lumbricoides*, *Trichuris trichiura* and hookworm infections into low, moderate and high

	A. lumbricoides	*T. trichiura*	Hookworm
Low	1−4999	1−999	1−1999
Moderate	5000−49,999	1000−9999	2000−3999
High	≥50,000	≥10,000	≥4000

The *TP* of moderate intensity *A. lumbricoides* infections in this population *j* can be written as,

$$\begin{aligned} E\left[TP_{j_{mod_a}}\right] &= P\left(5000 \geq X_{ij} \leq 49,999\right) \\ &= P\left(X_{ij} \leq 49,999\right) - P\left(X_{ij} \leq 4999\right) \\ &= F\left(49,999; \mu_j, k_j\right) - F\left(4999; \mu_j, k_j\right) \end{aligned} \tag{34}$$

The *TP* of high intensity *A. lumbricoides* infections in this population can be written as,

$$\begin{aligned} E\left[TP_{j_{high_a}}\right] &= P\left(X_{ij} \geq 50,000\right) \\ &= 1 - P\left(X_{ij} \leq 49,999\right) \\ &= 1 - F\left(49,999; \mu_j, k_j\right) \end{aligned} \tag{35}$$

In each of these formulae, F represents the cumulative distribution function (lower tail) of the negative binomial distribution with mean μ_j and aggregation parameter k_j. When individual stool samples are screened we replace μ_j by $\overline{Y_{\cdot j}}$ and k_j by $\overline{Y_{\cdot j}}/(Var[Y_{ij}]/\overline{Y_{\cdot j}} - 1/f_j - 1)$. When pooled samples are examined, we replace μ_j by $\overline{U_{\cdot j}}$ and k_j by $\overline{U_{\cdot j}}/(m_j \cdot (Var[U_{lj}]/\overline{U_{\cdot j}} - 1/f_j) - 1)$.

In the hypothetical example described in Section 4.3.1, estimation of the *TP* of individual STH species, the *TP* of low, moderate and high *A. lumbricoides* intensity infections equalled 54.9%, 2.4% and 0.0001%, respectively. By substituting expected values and variances in Eqns (33)–(35) by the observed means and variances, we obtain a *TP* (95%CI) of 62.2% (31.3; 93.6) (low intensity), 1.9% (0.4; 2.7) (moderate intensity) and 0.0001% (0; 0.0001) (high intensity) based on examination of individual stool samples, and a *TP* of 61.3% (28.4; 98.7) (low intensity), 1.9% (0.0001; 2.9) (moderate intensity) and 0.0001% (0; 0.0001) (high intensity) for the examination of pooled stool samples. The 95%CI were obtained by a bootstrap analysis which took into account the correlation between mean and variance separately for Y_{ij} and U_{lj}.

4.4 Estimating μ_j and k_j

For the different applications, it is essential to define the expected negative binomial distribution of the FECs by providing the expected values for μ_j and k_j for the population *j* in which the survey will be conducted. When detailed data are available, these parameters can be easily estimated based on the individual FEC data. However, when this level of information is not at hand, an alternative approach to estimate both parameters is required.

As illustrated in Figure 2, μ_j and k_j are correlated, and hence when prior knowledge is at hand for one variable, it is possible to estimate the other one. Because prior information on mean population FEC μ_j is generally at hand, we will estimate k_j as a function of μ_j.

To this end, we applied a linear model on the epidemiological survey conducted in Kenya to estimate the k_j in a school j based on its corresponding μ_j. We extracted the results obtained by duplicate Kato-Katz thick smear ($f_j = 2 \cdot 0.0417$ g). We used $\overline{Y_{\cdot j}}$ and $\overline{Y_{\cdot j}}/(Var[Y_{ij}]/\overline{Y_{\cdot j}} - 12 - 1)$ as a proxy of μ_j and k_j, respectively. Schools with μ_j equal to zero or $k_j < 0$ were omitted from the analysis. Figure 15 illustrates the predicted k_j based on μ_j for *A. lumbricoides*, *T. trichiura* and hookworm, respectively. The equations to estimate k_j as a function of μ_j are provided in Table 5. It is important to note that these equations obtained form a restrict range of μ_j, and hence any extrapolation from these equations beyond this range should be implemented with care.

5. CONCLUSIONS

In the present study, we described a general mathematical framework for egg counts in stool. Subsequently, we discussed how to estimate the sample size for assessing the population mean FEC and the impact of an intervention (measured as FECR) for any scenario of STH epidemiology (mean population FEC and aggregation of FEC between individuals) and diagnostic strategy (amount of stool examined ~ sensitivity of the diagnostic technique, and examination of individual/pooled samples). We illustrated how to calculate a sample size for (1) assessing population mean FEC with a predefined level of precision, (2) determining whether population mean FEC exceeds a predefined level of infection intensity, (3) correctly classifying a population free of STH with a predefined probability, (4) assessing impact of an intervention with a predefined level of precision and (5) examining whether the efficacy of a drug remains satisfactory, as well as estimating the *TP* of (6) each individual STH species separately, (7) any STH species and (8) low, moderate and high intensity infections in the absence of a gold standard.

Based on a prior knowledge on the local epidemiology, health decision-makers can now compare different diagnostic strategies and their corresponding technical and financial resources required, and hence optimize the use of funds allocated for monitoring MDA programmes to control STH. In the present study we focussed on STH infections, but this framework can be generalized to any other helminth infection, both in public or animal health, that is diagnosed by detection and quantification of eggs

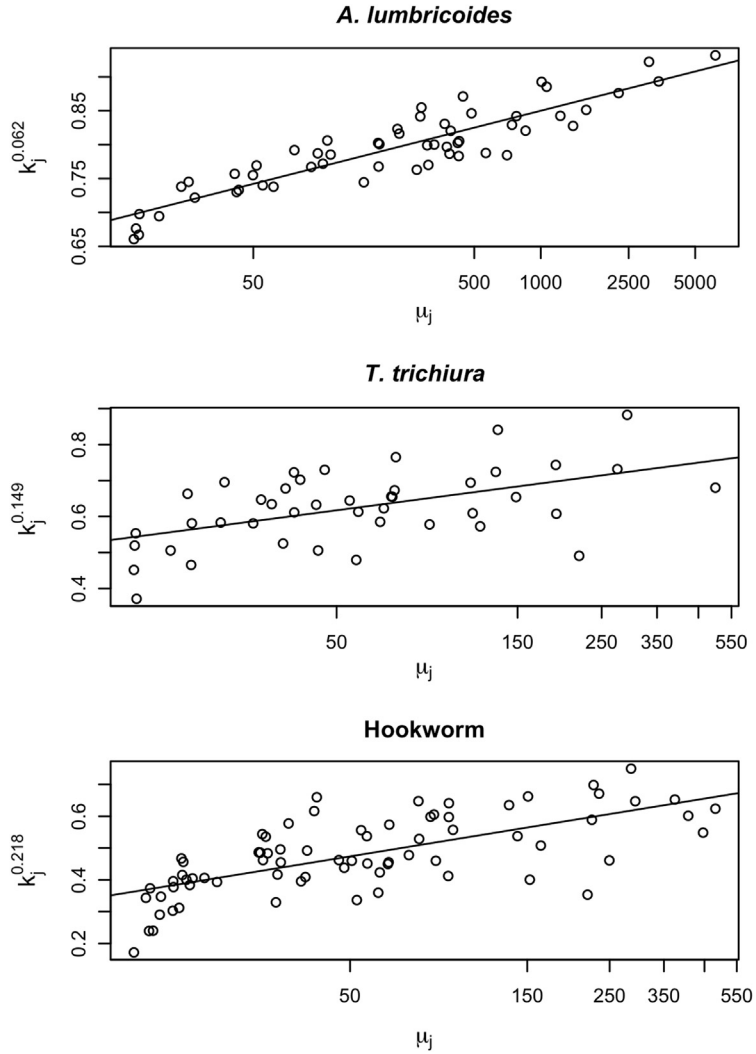

Figure 15 The scatter plots of aggregation parameter k_j as a function of mean school faecal egg counts (FECs; expressed in eggs per gram of stool (EPG)) μ_j, expressed in EPG of stool for *Ascaris lumbricoides*, *Trichuris trichiura* and hookworm. The dataset used was collected during an epidemiological survey in Kenya (Brooker et al., 2012).

in stool. For example, in public health the framework can be applied for *Schistosoma mansoni*, the causative agent of schistosomiasis (WHO, 2011). In animal health, we can apply the framework to gastrointestinal nematode infections, which up to date account for important losses in production of livestock (Charlier et al., 2014).

Table 5 Equations to estimate the aggregation parameter k_j based on the mean population faecal egg count (FEC; expressed in eggs per gram of stool (EPG)) μ_j, and the range of μ_j in which these equations were obtained for *Ascaris lumbricoides*, *Trichuris trichiura* and hookworm infections. The dataset used was collected during an epidemiological survey in Kenya (Brooker et al., 2012)

	***A. lumbricoides* $n = 56$**	***T. trichiura* $n = 43$**	**Hookworms $n = 73$**
k_j	$\sqrt[0.062]{0.060 + 0.036 \cdot \log(\mu_j)}$	$\sqrt[0.149]{0.381 + 0.060 \cdot \log(\mu_j)}$	$\sqrt[0.218]{0.150 + 0.083 \cdot \log(\mu_j)}$
min μ_j	14	14	13
max μ_j	6173	499	482

Although this study allows healthcare decision-makers to adapt their survey design according to both local helminth epidemiology and resources, there are some limitations that need to be acknowledged. First, it is important to note that this mathematical framework only holds true under the assumptions made, and hence that any sample size and *TP* obtained through the framework should be interpreted with caution. In the current framework we assumed that (1) individual FECs within a population follow a negative binomial distribution, (2) the number of eggs observed by microscopy in *f* g of stool follows a Poisson distribution, (3) the true FEC in a pooled stool sample equals the mean of the true FECs of the individual stool samples pooled and that (4) the impact of an intervention is equal for all subjects. As discussed in Section 2, sources of variability in egg counts, these assumptions explain a large proportion of the variation, but definitely not all of the variation observed in the field. For example, a zero-inflated negative binomial may be a more appropriate fit in scenarios where zero FECs are more frequent, and additional sources of variation will need to be parameterized to grasp all the variation observed in egg excretion (e.g. day-to-day variation), the egg counting process (e.g. variation in FEC between diagnostic techniques and laboratory technicians, and the impact of an intervention (see Figure 7). On the other hand, we should be aware that expanding the model by parameterizing the remaining variation, if possible at all, would compromise the accessibility of the framework for the end users: more parameters will need to be defined and prior information on these additional may not always be at hand. Second, if one wants to estimate the mean FEC in school children within a district by first sampling a number of schools followed by sampling a number of subjects per school (two-stage cluster sampling), one cannot use the framework to calculate the total number of children that needs to be screened over these schools. This is because the current model does not account for clustering of STH infections between populations, and hence ignores the additional variation in FEC between clusters (*in casu* schools). Given the additional complexity of clustered STH infections, we will work out an extended model in a follow-up study. Finally, it is obvious that the framework in its current form will not be attractive to a wide spectrum of possible end users. To bridge the gap between this study and the end users we plan to develop an online tool that provides sample size and *TP* estimates without the need of prior knowledge on the mathematical framework or any statistical software.

APPENDICES

A. The expected value and variance of faecal egg counts (FECs) based on individual stool samples

A.1 FECs prior to an intervention

A.1.1 The expected value

$$\begin{aligned} E\left[Y_{ij}\right] &= E\left[\frac{Z_{ij}}{f_j}\right] = \frac{E\left[Z_{ij}\right]}{f_j} \\ &= \frac{E\left[E\left[Z_{ij}\middle|\lambda_{ij}, X_{ij}\right]\right]}{f_j} \\ &= \frac{E\left[f_j \cdot X_{ij}\right]}{f_j} \\ &= \frac{f_j \cdot E\left[X_{ij}\right]}{f_j} \\ &= \mu_j \end{aligned}$$

A.1.2 The variance

$$\begin{aligned} Var\left[Y_{ij}\right] &= Var\left[\frac{Z_{ij}}{f_j}\right] = \frac{Var\left[Z_{ij}\right]}{f_j^2} \\ &= \frac{1}{f_j^2} \cdot \left(E\left[Var\left[Z_{ij}\middle|\lambda_{ij}, X_{ij}\right]\right] + Var\left[E\left[Z_{ij}\middle|\lambda_{ij}, X_{ij}\right]\right]\right) \\ &= \frac{1}{f_j^2} \cdot \left(E\left[f \cdot X_{ij}\right] + Var\left[f_j \cdot X_{ij}\right]\right) \\ &= \frac{1}{f_j^2} \cdot \left(f_j \cdot E\left[X_{ij}\right] + f_j^2 \cdot Var\left[X_{ij}\right]\right) \\ &= \frac{1}{f_j^2} \cdot \left(f_j \cdot \mu_j + f_j^2 \cdot \left(\mu_j + \mu_j^2 \middle/ k_j\right)\right) \\ &= \frac{f_j^2 \cdot \mu_j}{f_j^2} \cdot \left(\frac{1}{f_j} + 1 + \frac{\mu_j}{k_j}\right) \\ &= \mu_j \cdot \left(\frac{1}{f_j} + 1 + \frac{\mu_j}{k_j}\right) \end{aligned}$$

A.2 FECs after an Intervention

A.2.1 The expected value

$$
\begin{aligned}
E\left[Y'_{ij}\right] &= E\left[\frac{Z'_{ij}}{f_j}\right] = \frac{E\left[Z'_{ij}\right]}{f_j} \\
&= \frac{E\left[E\left[Z'_{ij}\middle|\lambda'_{ij}, X'_{ij}\right]\right]}{f_j} \\
&= \frac{E\left[f_j \cdot X'_{ij}\right]}{f_j} \\
&= \frac{f_j \cdot E\left[X'_{ij}\right]}{f_j} \\
&= E\left[\left(1-e_j\right)\cdot X_{ij}\right] \\
&= \left(1-e_j\right)\cdot E\left[X_{ij}\right] \\
&= \left(1-e_j\right)\cdot \mu_j
\end{aligned}
$$

A.2.2 The variance

$$
\begin{aligned}
Var\left[Y'_{ij}\right] &= Var\left[\frac{Z'_{ij}}{f_j}\right] = \frac{Var\left[Z'_{ij}\right]}{f_j^2} \\
&= \frac{1}{f_j^2}\cdot\left(E\left[Var\left[Z'_{ij}\middle|\lambda'_{ij}, X'_{ij}\right]\right] + Var\left[E\left[Z'_{ij}\middle|\lambda'_{ij}, X'_{ij}\right]\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(E\left[f_j \cdot X'_{ij}\right] + Var\left[f_j \cdot X'_{ij}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j \cdot E\left[X'_{ij}\right] + f_j^2 \cdot Var\left[X'_{ij}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j \cdot E\left[\left(1-e_j\right)\cdot X_{ij}\right] + f_j^2 \cdot Var\left[\left(1-e_j\right)\cdot X_{ij}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j \cdot \left(1-e_j\right)\cdot E\left[X_{ij}\right] + f_j^2 \cdot \left(1-e_j\right)^2 \cdot Var\left[X_{ij}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j \cdot \left(1-e_j\right)\cdot \mu_j + f_j^2 \cdot \left(1-e_j\right)^2 \cdot \left(\mu_j + \mu_j^2 \middle/ k_j\right)\right) \\
&= \frac{f_j^2 \cdot \left(1-e_j\right)\cdot \mu_j}{f_j^2}\cdot\left(\frac{1}{f_j} + \left(1-e_j\right)\cdot\left(1 + \mu_j \middle/ k_j\right)\right) \\
&= \left(1-e_j\right)\cdot \mu_j \cdot \left(\frac{1}{f_j} + \left(1-e_j\right)\cdot\left(1 + \mu_j \middle/ k_j\right)\right)
\end{aligned}
$$

B. The expected value and variance of FECs based on pools of m_j individual stool samples

B.1 FECs prior to an intervention

B.1.1 The expected value

$$
\begin{aligned}
E\left[U_{lj}\right] &= E\left[\frac{Z_{lj}}{f_j}\right] = \frac{E\left[Z_{lj}\right]}{f_j} \\
&= \frac{E\left[E\left[Z_{lj}\middle|\lambda_{lj}, V_{lj}\right]\right]}{f_j} \\
&= \frac{E\left[f_j \cdot Vlj\right]}{f_j} \\
&= \frac{f_j \cdot E\left[Vlj\right]}{f_j} \\
&= E\left[\frac{\sum_{i=1}^{m_j} X_{ij}}{m_j}\right] \\
&= \frac{m_j}{m_j} \cdot E\left[X_{ij}\right] \\
&= \mu_j
\end{aligned}
$$

B.1.2 The variance

$$
\begin{aligned}
Var\left[U_{lj}\right] &= Var\left[\frac{Z_{lj}}{f_j}\right] = \frac{Var\left[Z_{lj}\right]}{f_j^2} \\
&= \frac{1}{f_j^2}\cdot\left(E\left[Var\left[Z_{lj}\middle|\lambda_{lj}, V_{lj}\right]\right] + Var\left[E\left[Z_{lj}\middle|\lambda_{lj}, V_{lj}\right]\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(E\left[f_j\cdot V_{lj}\right] + Var\left[f_j\cdot V_{lj}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j\cdot E\left[V_{lj}\right] + f_j^2\cdot Var\left[V_{lj}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j\cdot E\left[\frac{\sum_{i=1}^{m_j} X_{ij}}{m_j}\right] + f_j^2\cdot Var\left[\frac{\sum_{i=1}^{m_j} X_{ij}}{m_j}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(\frac{f_j\cdot m_j}{m_j}\cdot E\left[X_{ij}\right] + \frac{f_j^2\cdot m_j}{m_j^2}\cdot Var\left[X_{ij}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j\cdot\mu_j + \frac{f_j^2}{m_j}\cdot\left(\mu_j + \mu_j^2/k_j\right)\right) \\
&= \frac{f_j^2\cdot\mu_j}{f_j^2}\cdot\left(\frac{1}{f_j} + \frac{1+\mu_j/k_j}{m_j}\right) \\
&= \mu_j\cdot\left(\frac{1}{f_j} + \frac{1+\mu_j/k_j}{m_j}\right)
\end{aligned}
$$

B.2 FECs after an intervention

B.2.1 The expected value

$$
\begin{aligned}
E\left[U'_{lj}\right] &= E\left[\frac{Z'_{lj}}{f_j}\right] = \frac{E\left[Z'_{lj}\right]}{f_j} \\
&= \frac{E\left[E\left[Z'_{lj}\middle|\lambda'_{lj}, V'_{lj}\right]\right]}{f_j} \\
&= \frac{E\left[f_j \cdot V'_{lj}\right]}{f_j} \\
&= \frac{f_j \cdot E\left[V'_{lj}\right]}{f_j} \\
&= E\left[\frac{\sum_{i=1}^{m_j} X'_{lj}}{m_j}\right] \\
&= \frac{m_j}{m_j} \cdot E\left[X'_{lj}\right] \\
&= E\left[\left(1 - e_j\right) \cdot X_{ij}\right] \\
&= \left(1 - e_j\right) \cdot E\left[X_{ij}\right] \\
&= \left(1 - e_j\right) \cdot \mu_j
\end{aligned}
$$

B.2.2 The variance

$$
\begin{aligned}
Var\left[U'_{lj}\right] &= Var\left[\frac{Z'_{lj}}{f_j}\right] = \frac{Var\left[Z'_{lj}\right]}{f_j^2} \\
&= \frac{1}{f_j^2}\cdot\left(E\left[Var\left[Z'_{lj}\middle|\lambda'_{lj}, V'_{lj}\right]\right] + Var\left[E\left[Z'_{lj}\middle|\lambda'_{lj}, V'_{lj}\right]\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(E\left[f_j\cdot V'_{lj}\right] + Var\left[f_j\cdot V'_{lj}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j\cdot E\left[V'_{lj}\right] + f_j^2\cdot Var\left[V'_{lj}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j\cdot E\left[\frac{\sum_{i=1}^{m_j} X'_{ij}}{m_j}\right] + f_j^2\cdot Var\left[\frac{\sum_{i=1}^{m_j} X'_{ij}}{m_j}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(\frac{f_j\cdot m_j}{m_j}\cdot E\left[X'_{ij}\right] + \frac{f_j^2\cdot m_j}{m_j^2}\cdot Var\left[X'_{ij}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j\cdot E\left[\left(1-e_j\right)\cdot X_{ij}\right] + \frac{f_j^2}{m_j}\cdot Var\left[\left(1-e_j\right)\cdot X_{ij}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j\cdot\left(1-e_j\right)\cdot E\left[X_{ij}\right] + \frac{f_j^2\cdot\left(1-e_j\right)^2}{m_j}\cdot Var\left[X_{ij}\right]\right) \\
&= \frac{1}{f_j^2}\cdot\left(f_j\cdot\left(1-e_j\right)\cdot\mu_j + \frac{f_j^2\cdot\left(1-e_j\right)^2}{m_j}\cdot\left(\mu_j + \mu_j^2\middle/k_j\right)\right) \\
&= \frac{f_j^2\cdot(1-e)\cdot\mu_j}{f_j^2}\cdot\left(\frac{1}{f_j} + \frac{1-e_j}{m_j}\cdot\left(1+\mu_j\middle/k_j\right)\right) \\
&= \left(1-e_j\right)\cdot\mu_j\cdot\left(\frac{1}{f_j} + (1-e)\cdot\frac{1+\mu_j\middle/k_j}{m_j}\right)
\end{aligned}
$$

ACKNOWLEDGMENTS

Bruno Levecke is a Postdoctoral fellow of FWO (www.fwo.be). The work of Johannes Charlier received funding from the E.U. FP7 GLOWORM project (Grant Agreement No. 288975CP-TP-KBBE.2011.1.3-04).

REFERENCES

Addiss, D.G., Rheingans, R., Twum-Danso, N.A., Richards, F.O., 2003. A framework for decision-making for mass distribution of Mectizan in areas endemic of *Loa loa*. Filaria J. 2 (Suppl. 1), S9.

Albonico, M., Rinaldi, L., Sciascia, S., Morgoglione, M.E., Piemonte, M., et al., 2013. Comparison of three copromicroscopic methods to assess albendazole efficacy against soil-transmitted helminth infections in school-aged children on Pemba Island. Trans. R. Soc. Trop. Med. Hyg. 107, 493–501.

Anderson, R., Hollingsworth, T.D., Truscott, J., Brooker, S.J., 2012. Optimisation of mass chemotherapy to control soil-transmitted helminth infection. Lancet 379, 289–290.

Barda, B., Zepherine, H., Rinaldi, L., Cringoli, G., Burioni, R., et al., 2013. Mini-FLOTAC and Kato-Katz: helminth eggs watching on the shore of Lake Victoria. Parasit. Vectors 6, 220.

Bergquist, R., Johansen, M.V., Utzinger, J., 2009. Diagnostic dilemmas in helminthology: what tools to use and when? Trends Parasitol. 25, 151–156.

Bethony, J., Brooker, S.J., Albonico, M., Geiger, S.M., Loukas, A., et al., 2006. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. Lancet 367, 1521–1532.

Bogoch, I.I., Raso, G., NGoran, E.K., Marti, H.P., Utzinger, J., 2006. Differences in microscopic diagnosis of helminths and intestinal protozoa among diagnostic centres. Eur. J. Clin. Microbiol. Infect. Dis. 25, 344–347.

Booth, M., Vounatsou, P., NGoran, E.K., Tanner, M., Utzinger, J., 2003. The influence of sampling effort and the performance of the Kato-Katz technique in diagnosing *Schistosoma mansoni* and hookworm co-infections in rural Côte d Ivoire. Parasitology 127, 525–531.

Brooker, S.J., Hotez, P.J., Bundy, D.A.P., 2010. The global atlas of helminth infection: mapping the way forward in neglected tropical disease control. PLoS Negl. Trop. Dis. 4, e779.

Brooker, S.J., Kabatereine, N.B., Myatt, M., Stothard, R.J., Fenwick, A., 2005. Rapid assessment of *Schistosoma mansoni*: the validity, applicability and cost-effectiveness of the lot quality assurance sampling method in Uganda. Trop. Med. Int. Health 10, 647–658.

Brooker, S.J., Pullan, R.L., Gitonga, C.W., Ashton, R.A., Kolaczinski, J.H., et al., 2012. *Plasmodium*–helminth coinfection and its sources of heterogeneity across East Africa. J. Infect. Dis. 205, 841–852.

Casella, G., Berger, B.L., 2002. Statistical Inference, second ed. Duxbury Press. Pacific Grove: Duxbury Thomson Learning.

Charlier, J., van der Voort, M., Kenyon, F., Skuce, P., Vercruysse, J., 2014. Chasing helminths and their economic impact on farmed ruminants. Trends Parasitol. 30, 361–367.

Cringoli, G., Rinaldi, L., Albonico, M., Bergquist, R., Utzinger, J., 2013. Geospatial (s)tools: integration of advanced epidemiological sampling and novel diagnostics. Geospat. Health 7, 399–404.

Cringoli, G., Rinaldi, L., Maurelli, M.P., Utzinger, J., 2010. FLOTAC: new multivalent techniques for qualitative and quantitative copromicroscopic diagnosis of parasites in animals and humans. Nat. Protoc. 5, 503–515.

de Silva, N., Hall, A., 2010. Using the prevalence of individual species of intestinal nematode worms to estimate the combined prevalence of any species. PLoS Negl. Trop. Dis. 4, e655.

Diggle, P.J., 2011. Estimating prevalence using an imperfect test. Epidemiol. Res. Int. 608719.

Dobson, R.J., Sangster, N.C., Besier, R.B., Woodgate, R.G., 2009. Geometric means provide a biased efficacy result when conducting a faecal egg count reduction test (FECRT). Vet. Parasitol. 161, 162–167.

Enœ, C., Georgiadis, M.P., Johnson, W.O., 2000. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. Prev. Vet. Med. 45, 6181.

Gregory, R.D., Woolhouse, M.E., 1993. Quantification of parasite aggregation: a simulation study. Acta Trop. 54, 131–139.

Hui, S.L., Zhou, X.H., 1998. Evaluation of diagnostic tests without gold standards. Stat. Methods Med. Res. 7, 354–370.

Hürlimann, E., Schur, N., Boutsika, K., Stensgaard, A.-S., Himpsl, Laserna de, et al., 2011. Toward an open-access global database for mapping, control, and surveillance of neglected tropical diseases. PLoS Negl. Trop. Dis. 5, e1404.

Humphries, D., Mosites, E., Otchere, J., Twum, W.A., Woo, L., et al., 2011. Epidemiology of hookworm infection in Kintampo North Municipality, Ghana: patterns of malaria coinfection, anemia, and albendazole treatment failure. Am. J. Trop. Med. Hyg. 84, 792–800.

Katz, N., Chaves, A., Pellegrino, J., 1972. A simple device for quantitative stool thick-smear technique in schistosomiasis mansoni. Rev. Inst. Med. Trop. São Paulo 14, 397–400.

Knopp, S., Glinz, D., Rinaldi, L., Mohammed, K.A., NGoran, E.K., et al., 2009. FLOTAC: a promising technique for detecting helminth eggs in human feces. Trans. R. Soc. Trop. Med. Hyg. 103, 1190–1194.

Knopp, S., Mgeni, A.F., Khamis, I.S., Steinmann, P., Stothard, J.R., et al., 2008. Diagnosis of soil-transmitted helminths in the era of preventive chemotherapy: effect of multiple stool sampling and use of different diagnostic techniques. PLoS Negl. Trop. Dis. 2, e331.

Krauth, S.J., Coulibaly, J.T., Knopp, S., Traoré, M., NGoran, E.K., et al., 2012. An in-depth analysis of a piece of shit: distribution of *Schistosoma mansoni* and hookworm eggs in human stool. PLoS Negl. Trop. Dis. 6, e1969.

Levecke, B., Behnke, J.M., Ajjampur, S.S., Albonico, M., Ame, S.M., et al., 2011b. A comparison of the sensitivity and fecal egg counts of the McMaster egg counting and Kato-Katz thick smear methods for soil-transmitted helminths. PLoS Negl. Trop. Dis. 5, e1201.

Levecke, B., Brooker, S.J., Knopp, S., Steinmann, P., Stothard, R.J., et al., 2014a. Effect of sampling and diagnostic effort on the assessment of schistosomiasis and soil-transmitted helminthiasis and drug efficacy: a meta-analysis of six drug efficacy trials and one epidemiological survey. Parasitology 141, 1826–1840.

Levecke, B., Speybroeck, N., Dobson, R., Vercruysse, J., Charlier, J., 2011a. Novel insights in the fecal egg count reduction test for monitoring drug efficacy against soil-transmitted helminths in large-scale treatment programs. PLoS Negl. Trop. Dis. 5, e1427.

Levecke, B., Montresor, A., Albonico, M., Ame, S.M., Behnke, J.M., et al., 2014b. Anthelminthic efficacy of single-oral dose 500 mg mebendazole in school children in six countries where soil-transmitted helminths are endemic: a meta-analysis. PLoS Negl. Trop. Dis. 8, e3204.

Ministry of Agriculture, Fisheries and Food, 1986. Manual of Veterinary Parasitological Laboratory Techniques (Reference Book; 418), third ed. Her Majesty's Stationery Office, London.

Montresor, A., 2011. Cure rate is not a valid indicator for assessing drug efficacy and impact of preventive chemotherapy interventions against schistosomiasis and soil-transmitted helminthiasis. Trans. R. Soc. Trop. Med. Hyg. 105, 361–363.

Montresor, A., Engels, D., Chitsulo, L., Gabrielli, A., Albonico, M., et al., 2011. The appropriate indicator should be used to assess treatment failure in STH infections. Am. J. Trop. Med. Hyg. 85, 579–580.

Morgan, E.R., Cavill, L., Curry, G.E., Wood, R.M., Mitchell, E.S., 2005. Effects of aggregation and sample size on composite faecal egg counts in sheep. Vet. Parasitol. 131, 79–87.

Murray, C.J.L., Vos, T., Lozano, R., Naghavi, M., Flaxman, A.D., et al., 2012. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 19902010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 9859, 2179–2223.

Neglected Tropical Disease Partner Website, 2013. Uniting to Combat Neglected Tropical Diseases. Ending the Neglect and Reaching 2020 Goals. Available at: http://www.unitingtocombatntds.org (accessed 10.04.14.).

Olives, C., Valadez, J.J., Brooker, S.J., Pagano, M., 2012. Multiple category-lot quality assurance sampling: a new classification system with application to schistosomiasis control. PLoS Negl. Trop. Dis. 6, e1806.

Pullan, R.L., Brooker, S.J., 2012. The global limits and population at risk of soil-transmitted helminth infections in 2010. Parasit. Vectors 5, 81.

Pullan, R.L., Smith, J.L., Jasrasaria, R., Brooker, S.J., 2014. Global numbers of infection and disease burden of soil-transmitted helminth infections in 2010. Parasit. Vectors 7, 37.

Sinniah, B., 1982. Daily egg production of *Ascaris lumbricoides*: the distribution of eggs in the feces and the variability of egg counts. Parasitology 84, 167–175.

Speybroeck, N., Devleesschauwer, B., Joseph, L., Berkvens, D., 2013. Misclassification errors in prevalence estimation: Bayesian handling with care. Int. J. Public Health 58, 791–795.

Torgerson, P.R., Paul, M., Lewis, F.I., 2012. The contribution of simple random sampling to observed variations in faecal egg counts. Vet. Parasitol. 188, 397–401.

Torgerson, P.R., Schnyder, M., Hertzberg, H., 2005. Detection of anthelmintic resistance: a comparison of mathematical techniques. Vet. Parasitol. 128, 291–298.

Vercruysse, J., Behnke, J.M., Albonico, M., Ame, S.M., Angebault, C., et al., 2011. Assessment of anthelminthic efficacy of albendazole in school children in seven countries where soil-transmitted helminths are endemic. PloS Negl. Trop. Dis. 5, e948.

World Health Organization, 1991. Basic Laboratory Methods in Medical Parasitology. World Health Organization, Geneva, Switzerland.

World Health Organization, 1998. Guidelines for the Evaluation of Soil-transmitted Helminthiasis and Schistosomiasis at Community Level. A Guide for Control Programme Managers. World Health Organization, Geneva, Switzerland.

World Health Organization, 2011. Helminth Control in School-age Children: A Guide for Managers of Control Programmes, Second ed. World Health Organization, Geneva, Switzerland.

World Health Organization, 2012a. Accelerating Work to Overcome the Global Impact of Neglected Tropical Diseases: A Roadmap for Implementation. World Health Organization, Geneva, Switzerland.

World Health Organization, 2012b. Soil-transmitted helminthiases: number of children treated in 2012. Week. Epidemiol. Rec. 89, 133–140.

World Health Organization, 2012c. Soil-transmitted Helminthiases: Eliminating Soil-transmitted Helminthiases as a Public Health Problem in Children: Progress Report 2001–2010 and Strategic Plan 2011–2020. World Health Organization, Geneva, Switzerland.

World Health Organization, 2013. Assessing the Efficacy of Anthelminthic Drugs against Schistosomiasis and Soil-Transmitted Helminthiasis. World Health Organization, Geneva, Switzerland.

Ye, X.P., Donnelly, C.A., Fu, Y.L., Wu, Z.X., 1997. The non-randomness of the distribution of *Trichuris trichiura* and *Ascaris lumbricoides* in feces and the effect of stirring fecal specimens. Trop. Med. Int. Health 2, 261–264.