ORIGINAL PAPER

Comparison of calculation methods used for the determination of anthelmintic resistance in sheep in a temperate continental climate

**L. C. Falzon · J. van Leeuwen · P. I. Menzies ·
A. Jones-Bitton · W. Sears · J. T. Jansen ·
A. S. Peregrine**

Abstract This study compared results obtained with five different fecal egg count reduction (FECR) calculation methods for defining resistance to ivermectin, fenbendazole, and levamisole in gastrointestinal nematodes of sheep in a temperate continental climate: $FECR_1$ and $FECR_2$ used pre- and posttreatment fecal egg count (FEC) means from both treated and control animals, but $FECR_1$ used arithmetic means, whereas $FECR_2$ used geometric means; $FECR_3$ used arithmetic means for pre- and posttreatment FECs from treated animals only; $FECR_4$ was calculated using only arithmetic means for posttreatment FECs from treated and control animals; and $FECR_5$ was calculated using mean FEC estimates from a general linear mixed model. The classification of farm anthelmintic resistance (AR) status varied, depending on which FECR calculation method was used and whether a bias correction term (BCT, i.e., half the minimum detection limit) was added to the zeroes or not. Overall, agreement between all methods was higher when a BCT was used, particularly when levels of resistance were low. $FECR_4$ showed the highest agreement with all the other FECR methods. We therefore recommend that small ruminant clinicians use the $FECR_4$ formula with a BCT for AR determination, as this would reduce the cost of the FECRT, while still minimizing bias and allowing for comparisons between different farms. For researchers, we recommend the use of $FECR_1$ or $FECR_2$, as the inclusion of both pre- and posttreatment FECs and use of randomly allocated animals in treatment and control groups makes these methods mathematically more likely to estimate the true anthelmintic efficacy.

Keywords Gastrointestinal nematodes · Fecal egg count reduction calculation methods · Arithmetic mean · Geometric mean · Bias correction term · Kappa

L. C. Falzon (✉) · P. I. Menzies · A. Jones-Bitton · W. Sears
Department of Population Medicine, Ontario Veterinary College,
University of Guelph, Guelph, ON N1G 2 W1, Canada
e-mail: laura.falzon@vetsuisse.unibe.ch

J. van Leeuwen
Centre for Veterinary Epidemiological Research, Department of
Health Management, Atlantic Veterinary College, University of
Prince Edward Island, Charlottetown, PE C1A 4P3, Canada

J. T. Jansen
Veterinary Science and Policy, Ontario Ministry of Agriculture and
Food, Elora, ON N0B 1S0, Canada

A. S. Peregrine
Department of Pathobiology, Ontario Veterinary College, University
of Guelph, Guelph, Ontario N1G 2 W1, Canada

Present Address:
L. C. Falzon
Veterinary Public Health Institute, Vetsuisse Faculty, University of
Bern, Schwarzenburgstrasse 155, 3097 Liebefeld, Switzerland

Introduction

Gastrointestinal nematodes (GINs) are of concern on sheep farms worldwide as they impair milk, meat, and wool production in sheep and are an important cause of morbidity and mortality (Knox et al. 2012). For many years, producers have relied primarily on the use of anthelmintics for the control of GINs in sheep (Sargison 2008). However, this reliance has led to the development of anthelmintic resistance (AR) and many countries are now reporting both multidrug and multinematode species resistance (Jackson and Coop 2000; Kaplan 2004; Kaplan and Vidyashankar 2012).

In many parasite populations of domesticated animals, a number of genotypically resistant parasites are present (Prichard et al. 1980). When these parasite populations exceed a certain frequency threshold, the parasite populations as a

whole become phenotypically resistant in animals (Kaplan and Vidyashankar 2012) and are associated with treatment failure and, inevitably, losses in sheep health and productivity (Kaplan 2004). It is therefore important to diagnose AR before it reaches this "critical frequency," and becomes a clinical and economic problem (Kaplan and Vidyashankar 2012).

The fecal egg count reduction test (FECRT) is the standard field test for the diagnosis of AR in sheep (Coles et al. 2006). However, the results from a FECRT may be influenced by several factors, including the study design, the host–parasite–environment interaction (Levecke et al. 2012a), and the mathematical formula used to calculate drug efficacy (Miller et al. 2006). The literature describes several methods to calculate the fecal egg count reduction (FECR); these methods differ based on whether the arithmetic or geometric mean is used, whether control animals are tested, and whether pre- and posttreatment fecal egg counts (FECs) are used in the calculation (Cabaret and Berrag 2004). Some authors have suggested the use of arithmetic means, as they are unbiased estimators of the true mean (Fulford 1994), and therefore provide better estimates of the parasite egg output compared with geometric means (Coles et al. 1992; Dobson et al. 2009). Meanwhile, other authors have described geometric means as more appropriate estimators of the central tendency parameter for parasite populations which are usually overdispersed and, therefore, do not have a constant variance (Smothers et al. 1999). While some FECR formulae take into account both pre- and posttreatment FECs in both treated and control animals (Presidente 1985; Dash et al. 1988), other variations of the FECR calculation only take into consideration the pre- and posttreatment FECs of treated animals (McKenna 2006), or the posttreatment FECs of both treated and untreated animals (Coles et al. 2006), thereby reducing the number of fecal samples required. Lastly, Mejia et al. (2003) described an alternate approach for calculating the FECR, using a general linear mixed model (GLMM) to provide FEC means corrected for other covariable effects, such as animal weight and treatment. It is therefore important to further assess the agreement between the classification of the farm resistance status using these different methods in order to improve and standardize the method of FECR calculation and limit misclassification of farm resistance status (Coles et al. 2006; Denwood et al. 2010).

Host–parasite–environment interactions can also influence FECRT results, as shown by simulation studies based on data from Belgium (Levecke et al. 2012a). In addition, Torgerson et al. (2005) examined AR to mebendazole, ivermectin, and moxidectin in a goat herd in Switzerland using both FEC and slaughter examination of worm burdens posttreatment, and found that commonly used mathematical techniques failed to detect low anthelmintic efficacy. However, the single herd and simulation models used in those European studies may not be representative of the general population of commercial sheep

and goats elsewhere, such as in temperate continental climates where different FECRT calculation methods for defining AR have not been previously examined.

The objective of this study was to compare the FECR percentages and AR classifications obtained using five different formulae, for resistance to ivermectin, fenbendazole, and levamisole among a typical group of commercial farms in a temperate continental climate.

Materials and methods

Farm selection, drench check, and fecal egg count reduction test

Full details of the farm selection and FECRT have been described in Falzon et al. (2013) and are summarized in Fig. 1. In brief, 47 sheep flocks across south-western Ontario, Canada, were enrolled over 2 years (in 2010 and 2011) in a study to determine the frequency of AR in Ontario sheep flocks. To be included in the study, farms had to: (i) have a minimum of 30 eligible animals in their first grazing season and (ii) keep the eligible animals on pasture for at least 3 months during the grazing season. Eligible animals included lambs (<6 months) or yearling ewes.

Producers participating in the study were asked to submit fecal samples from 15 eligible animals, and these were processed using a modified McMaster technique described by the Ministry of Agriculture, Fisheries and Food (1986) with a minimum limit of detection of 50 GIN eggs per gram (epg). If the mean FEC was ≥200 epg, the producers were asked to treat their eligible animals with ivermectin drench provided by the researchers; no instructions on how to treat or what dosage to use were provided, with the purpose that producers would treat as they customarily did. Fourteen days after treatment, the producers collected fecal samples from 15 of the treated eligible animals used in this drench check procedure and sent them to the laboratory for processing. The reduction in FECs was determined as: $100 \times ((\text{arithmetic mean FEC before treatment} - \text{arithmetic mean FEC after treatment}) / \text{arithmetic mean FEC before treatment})$; a reduction of <95 % was considered drench failure and suggestive of AR.

A FECRT for ivermectin (0.2 mg/kg), fenbendazole (5.0 mg/kg), and levamisole (10.5 mg/kg) was conducted by the research team on 29, 20, and 17 farms that had ivermectin drench failure, respectively. Fenbendazole and levamisole were tested on fewer farms due to a limited number of eligible animals on some farms meeting the inclusion criteria. A mean of 28 days (range=21 to 35 days) elapsed between ivermectin treatment for the drench check and the first visit for the FECRT. Ten to fifteen eligible animals (depending on the number available) were included in each treatment group, and fecal samples were collected on the day of treatment and

Fig. 1 A schematic diagram of the study design to determine the prevalence of anthelmintic resistance in south-western Ontario sheep flocks in 2010–2011 (as described in Falzon et al. 2013)
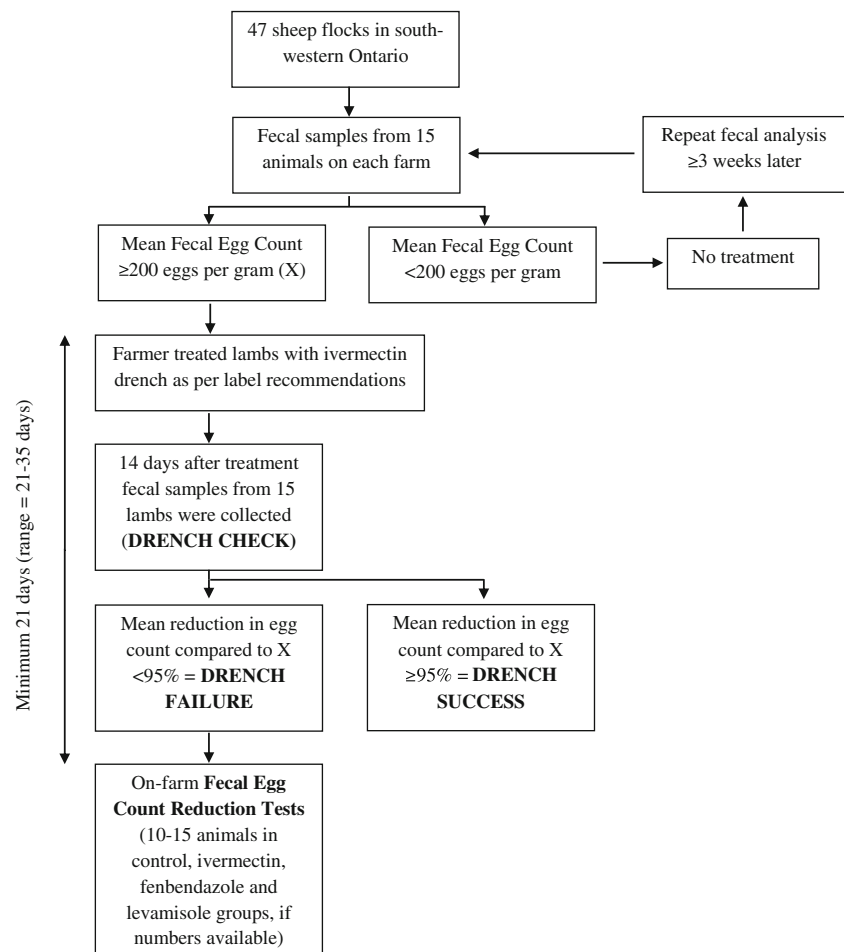
14 days later. On each farm, 10–15 eligible animals were also left untreated and acted as a control group. All eligible animals enrolled in the study were grazed together. This study was approved by the Animal Care Committee (protocol number 09R056) and the Research Ethics Board (protocol number 09DC005) at the University of Guelph.

Fecal egg count reduction calculations

Five different FECR calculations were conducted using SAS® 9.3 (SAS Institute Inc., Cary, NC, USA). As three calculations ($FECR_1$, $FECR_2$, and $FECR_5$) required the addition of a bias correction term (BCT) to allow for the computation of the mean and/or confidence intervals (CIs) (due to the presence of zero values), a BCT was used in all five methods to facilitate the comparison of results. Specifically, a 25 was added to all the zeroes; this value is half of the minimum detection limit of the diagnostic test used in this study (50 epg). The FECR calculations were computed as follows:

$$FECR_1 = 100 \times \left(1 - \left[\frac{T2}{T1}\right]\left[\frac{C1}{C2}\right]\right)$$

where $T1$ and $T2$ were pre- and posttreatment *arithmetic* means of the GIN epg in treated groups, respectively, and $C1$ and $C2$ were pre- and posttreatment *arithmetic* means of the epg in the controls (i.e., untreated animals), respectively (Dash et al. 1988).

$$FECR_2 = 100 \times \left(1 - \left[\frac{T2}{T1}\right]\left[\frac{C1}{C2}\right]\right)$$

where $T1$ and $T2$ were pre- and posttreatment *geometric* means of the epg in treated groups, respectively, and $C1$ and $C2$ were pre- and posttreatment *geometric* means of the epg in the controls, respectively (Presidente 1985).

For both $FECR_1$ and $FECR_2$, the 95 % CIs were estimated following the general approach described for standard errors of odds ratios and risk ratios (Kahn and Sempos 1989), while adjusting for the possible correlation between fecal samples collected before and after treatment from the same animal (Mood et al. 1974). A BCT was required for the computation of the CIs, as

variances and ratios are undefined when there are zeroes in the data, as shown in the following formulae used.

$$\text{Lower } 95\% \text{ limit} = \left(1 - \text{FECR}^{\left[\sqrt{\{\text{COVAR\_C} + \text{COVAR\_T}\}} \times 1.96\right]}\right) \times 100$$

$$\text{Upper } 95\% \text{ limit} = \left(1 - \text{FECR}^{\left[\sqrt{\{\text{COVAR\_C} + \text{COVAR\_T}\}} \times -1.96\right]}\right) \times 100$$

where COVAR_C (i.e., covariance between the pre- and posttreatment FECs of untreated sheep on a given farm) was calculated as:

$$\text{COVAR\_C} = \frac{1}{(C1 + 25)} + \frac{1}{(C2 + 25)} - \frac{2 \times R\_C}{\sqrt{([C1 + 25] \times [C2 + 25])}}$$

and COVAR_T (i.e., covariance between T1 and T2 on a given farm) was calculated as:

$$\text{COVAR}_T = \frac{1}{(T1 + 25)} + \frac{1}{(T2 + 25)} - \frac{2 \times R\_T}{\sqrt{([T1 + 25] \times [T2 + 25])}}$$

where R_C and R_T were the correlations between pre- and posttreatment FECs of the untreated and treated animals, respectively.

$$\text{SD\_FECR} = \frac{\sqrt{\left(\left[\dfrac{\text{SD\_}T2 \times 2}{T2}\right] + \left[\dfrac{(\text{SD\_}T1 \times 2)}{T1}\right] - \left[\dfrac{2 \times R_T \times \text{SD\_}T2 \times \text{SD\_}T1}{\dfrac{T2}{T1}}\right]\right)}}{N\_T2}$$

where SD_T1 and SD_T2 were the standard deviations of the *arithmetic* mean pre- and posttreatment FECs, respectively, and R_T was the correlation between pre- and posttreatment FECs. N_T2 was the number of animals in the posttreatment group.

$$\text{FECR}_4 = 100 \times \left(1 - \left[\frac{T2}{C2}\right]\right)$$

where T2 and C2 were the posttreatment *arithmetic* means of epg in the treated and control groups, respectively. Ninety-five percent CIs were estimated as:

$$95\% \text{ CI} = 100 \times \left(1 - \left[\frac{T2}{C2}\right]^{\{\mp 1.96\sqrt{Y^2}\}}\right)$$

where Y^2 was the variance of the difference between T2 and C2 (Coles et al. 1992).

$$\text{FECR}_3 = 100 \times \left(1 - \left[\frac{T2}{T1}\right]\right)$$

where T1 and T2 were pre- and posttreatment arithmetic means of the GIN epg in treated groups (McKenna 2006).

The 95 % CIs were calculated using the general approach for simple ratios (Kahn and Sempos 1989), while adjusting for the correlation between pre- and posttreatment fecal samples from the same animal (Mood et al. 1974):

$$\text{Lower } 95\% \text{ limit} = 100 \times \left(1 - \frac{T2^{[1.96 \times \text{SD\_FECR}]}}{T1}\right)$$

$$\text{Upper } 95\% \text{ limit} = 100 \times \left(1 - \frac{T2^{[-1.96 \times \text{SD\_FECR}]}}{T1}\right)$$

where SD_FECR (i.e., the standard deviation of the FECR percentage) was calculated as:

i. FECR$_5$ was calculated by building a GLMM, as described by Mejia et al. (2003). The GLMM was fit using PROC MIXED (SAS 9.3®), with the natural logarithm of the posttreatment FEC as the response variable. The dependence of the data was modeled by a fixed effect for farm (to obtain a coefficient for each farm) and a random slope parameter at the treatment level (due to different variances between treatment groups) (Littell et al. 1996). While accounting for the treatment random effect, each fixed effect variable was examined on its own to screen for variables to start the modeling process. Screened variables included the natural logarithm of the pretreatment FEC, animal weight, farm, treatment (control, ivermectin, fenbendazole, and levamisole), and year of study (first and second). Given the relatively small sample size, a liberal alpha value of ≤0.20 was used to indicate which terms to initially include in the model. The linearity of continuous variables was assessed graphically by plotting

lowess smoother curves and by including a quadratic term in the model, as described by Dohoo et al. (2009).

A final GLMM was built using a manual backwards step-wise procedure, by first including all variables that were significant in the univariable analyses ($p \leq 0.20$). After the main effects model was built ($p \leq 0.05$), predictors of interest that were not significant in the univariable analysis were forced into the model to assess potential confounding and conditional effects. All possible two-way interactions between significant predictors were tested. The model assumptions were assessed by plotting residuals against the predicted out-comes and explanatory variables, to look for homoscedastic-ity, nonlinearity, and outliers. Normality was visually assessed with histograms of the residuals and normal quantile plots and assessed statistically using four different tests offered by SAS (Shapiro–Wilk, Kolmogorov–Smirnov, Cramer–von Mises, and Anderson–Darling). Observations that were identified as outliers or influential were cross-checked with the original data sheets for any abnormality in the data to explain their behavior. The model was repeated without those observations, and differences in model estimates were noted. The predicted posttreatment FECs were back-transformed using the expo-nential function (equivalent to *geometric* means), and the FEC means were then used to calculate the FECR following the same method described for FECR$_4$ (Coles et al. 1992).

For all treatment groups, and for all FECR calculation methods, farms were classified as resistant when the reduction was <95 % and the lower 95 % CI was <90 %; if only one of these two criteria was met, the farm was classified as being suspected of resistance (Coles et al. 1992).

Comparison of data from fecal egg count reduction calculation methods

To calculate the Kappa values between the five different FECR methods, results were first categorized into ordinal variables (0=susceptible; 1=suspected resistance; 2=resis-tance). A weighted Kappa was calculated using Stata 12.1® (StataCorp LP, Texas, USA); the weighted Kappa better ac-counts for partial agreement between tests, compared with the nonweighted Kappa (Dohoo et al. 2009). Kappa values were then interpreted using the scale described by Dohoo et al. (2009), where values ≤ 0.0, or between 0.01 and 0.20, 0.21 and 0.40, 0.41 and 0.60, 0.61 and 0.80, and 0.81 and 1.00, were considered indicative of poor, slight, fair, moderate, substantial, and almost perfect agreement, respectively.

To determine whether the addition of a BCT was influential on the FECR percentages, we calculated FECR$_3$ and FECR$_4$ (which did not require a BCT for computation of the mean and/or CIs) without addition of the BCT, and re-computed the Kappa values between the calculated methods.

Table 1 summarizes all the abbreviations used in the paper.

Table 1 List of the abbreviations used in the paper

Abbreviation	Full term
AR	Anthelmintic resistance
BCT	Bias correction term
C1	Mean of gastrointestinal nematode egg counts in control (i.e., untreated) group at day 0
C2	Mean of gastrointestinal nematode egg counts in control (i.e., untreated) group at day 14
CI	Confidence interval
COVAR_C	Covariance between the pre- and posttreatment fecal egg counts of untreated sheep on a given farm
COVAR_T	Covariance between the pre- and posttreatment fecal egg counts of treated sheep on a given farm
EXP	Exponentiated
FEC	Fecal egg count
FECR	Fecal egg count reduction
FECRT	Fecal egg count reduction test
GIN	Gastrointestinal nematode
GLMM	Generalized linear mixed model
N_T2	Number of animals in the posttreatment group
R_C	Correlation between pre- and posttreatment fecal egg counts from the untreated animals
R_T	Correlation between pre- and posttreatment fecal egg counts from the treated animals
SD_FECR	Standard deviation of the fecal egg count reduction percentage
SD_T1	Standard deviation of the mean pretreatment fecal egg counts
SD_T2	Standard deviation of the mean posttreatment fecal egg counts
T1	Mean of gastrointestinal nematode egg counts in treated group at day 0 (pretreatment)
T2	Mean of gastrointestinal nematode egg counts in treated group at day 14 (posttreatment)
WAAVP	World Association for the Advancement of Veterinary Parasitology
Y^2	Variance of the difference between mean posttreatment FECs and mean control group FECs at day 14

Results

Descriptive results of different FECR calculation methods

Results of the ivermectin, fenbendazole, and levamisole FECR percentages and 95 % CIs, using the five different FECR methods are shown in Tables 2, 3, and 4, respectively. The FECR$_5$ percentage could not be calculated for two farms (farms 9 and 28), as no animal weight data were available for the control animals, and the weight variable was required in the GLMM estimate. Among the different calculation methods, the percentage of farms classified as resistant to ivermectin, fenbendazole, and levamisole ranged between 93 and 97 %, 88 and 100 %, and 24 and 35 %, respectively.

Table 2 Fecal egg count reduction (FECR) percentages (with 95 % confidence intervals) and anthelmintic resistance status, following ivermectin treatment on 29 sheep farms in south-western Ontario, Canada (2010 and 2011) calculated using five different FECR formulae

Farm	$FECR_1$		$FECR_2$		$FECR_3$		$FECR_4$		$FECR_5$	
1	37.5 (27.4 to 46.3)	R	−53.7 (−81.0 to −30.4)	R	73.5 (60.6 to 82.2)	R	76.1 (−13.0 to 95.0)	R	62.7 (62.5 to 62.9)	R
2	60.4 (52.2 to 67.3)	R	85.1 (82.4 to 87.4)	R	70.4 (17.9 to 89.3)	R	28.0 (−97.5 to 73.8)	R	**91.2 (91.1 to 91.3)**	**SR**
3	46.0 (−12.9 to 74.2)	R	62.5 (29.4 to 80.1)	R	−26.9 (−1189.1 to 87.5)	R	−93.2 (−1076.9 to 68.3)	R	23.6 (22.6 to 24.6)	R
4	28.5 (17.5 to 38.1)	R	45.4 (32.7 to 55.7)	R	39.1 (0.1 to 62.8)	R	28.7 (−42.1 to 64.3)	R	41.2 (41.1 to 41.3)	R
5	−25.7 (−58.8 to 0.6)	R	19.3 (−4.5 to 37.8)	R	25.7 (−0.9 to 45.3)	R	−8.6 (−87.8 to 37.2)	R	−15.2 (−15.3 to −15.0)	R
6	72.5 (65.5 to 78.0)	R	−8.7 (−38.9 to 15.0)	R	−21.1 (−105.8 to 28.7)	R	59.2 (26.3 to 77.4)	R	−40.1 (−40.2 to −40.0)	R
7	11.3 (−2.3 to 23.0)	R	−7.0 (−21.8 to 6.0)	R	−21.2 (−58.6 to 7.4)	R	−155.2(−565.4 to 2.1)	R	−1034.2 (−1041.1 to −1027.4)	R
8	−20.5 (−32.7 to −9.5)	R	−3.5 (−14.9 to 6.7)	R	−41.8 (−131.3 to 13.0)	R	−23.0 (−174.9 to 45.0)	R	52.8 (52.7 to 52.9)	R
9	18.8 (−4.7 to 37.1)	R	18.1 (−7.9 to 37.8)	R	−22.3 (−124.0 to 33.2)	R	53.9 (−3.0 to 79.4)	R	n/a	
10	83.1 (66.7 to 91.5)	R	52.9 (−3.5 to 78.5)	R	65.3 (48.2 to 76.7)	R	90.9 (52.2 to 98.3)	R	54.2 (53.6 to 54.9)	R
11	63.5 (54.1 to 71.0)	R	77.9 (73.4 to 81.7)	R	61.7 (30.1 to 79.0)	R	74.8 (48.4 to 87.7)	R	84.7 (84.6 to 84.8)	R
12	47.0 (38.2 to 54.6)	R	41.1 (35.1 to 46.5)	R	33.0 (−21.4 to 63.1)	R	19.6 (−85.2 to 65.1)	R	46.6 (46.5 to 46.7)	R
13	−0.5 (−33.3 to 24.3)	R	64.7 (55.4 to 72.1)	R	39.6 (−44.8 to 74.8)	R	56.4 (−26.5 to 85.0)	R	68.6 (68.5 to 68.7)	R
14	45.7 (34.7 to 54.9)	R	30.9 (15.5 to 43.6)	R	71.7 (30.9 to 88.4)	R	54.5 (−71.8 to 87.9)	R	42.4 (42.2 to 42.7)	R
15	−7.0 (−22.7 to 6.6)	R	33.0 (25.5 to 39.8)	R	41.0 (19.6 to 56.6)	R	−25.6 (−173.9 to 42.4)	R	56.6 (56.5 to 56.7)	R
16	**97.7 (94.8 to 99.1)**	**S**	**96.0 (90.5 to 98.3)**	**S**	**96.0 (89.6 to 98.5)**	**SR**	**95.1 (86.5 to 98.2)**	**SR**	**93.7 (93.6 to 93.8)**	**SR**
17	7.7 (−88.1 to 54.7)	R	−2.4 (−110.8 to 50.2)	R	−804.3 (−1849.3 to −319.6)	R	−22.7 (−110.0 to 28.1)	R	−43.1 (−43.2 to −43.1)	R
18	61.3 (56.9 to 65.2)	R	75.8 (73.1 to 78.2)	R	−40.2 (−168.9 to 26.9)	R	54.1 (−54.3 to 86.3)	R	68.1 (68.0 to 68.2)	R
19	39.8 (32.5 to 46.3)	R	13.0 (7.4 to 18.2)	R	3.5 (−106.5 to 54.9)	R	−4.3 (−126.6 to 52.0)	R	−27.2 (−27.3 to −27.1)	R
20	44.1 (33.9 to 52.7)	R	47.2 (42.0 to 51.9)	R	−45.0 (−133.0 to 9.8)	R	46.4 (−104.8 to 86.0)	R	65.7 (65.6 to 65.8)	R
21	84.8 (80.2 to 88.3)	R	61.7 (50.0 to 70.7)	R	79.4 (56.6 to 90.2)	R	66.8 (−8.9 to 89.9)	R	−2.6 (−3.8 to −1.4)	R
22	57.2 (52.1 to 61.74)	R	56.1 (49.0 to 62.3)	R	18.3 (−22.4 to 45.5)	R	69.4 (−33.7 to 93.0)	R	58.5 (58.1 to 58.9)	R
23	81.5 (76.0 to 85.7)	R	60.2 (51.3 to 67.4)	R	57.6 (13.0 to 79.3)	R	80.4 (58.4 to 90.8)	R	67.1 (67.0 to 67.2)	R
24	55.8 (36.9 to 69.0)	R	36.9 (16.1 to 52.6)	R	14.8 (−18.7 to 38.9)	R	−7.5 (−177.4 to 58.3)	R	−13.4 (−14.2 to 12.6)	R
25	86.9 (48.7 to 96.7)	R	45.9 (−95.7 to 85.1)	R	**95.9 (81.3 to 99.1)**	**SR**	61.2 (1.8 to 84.7)	R	20.5 (18.9 to 22.0)	R
26	78.8 (51.7 to 90.7)	R	−23.4 (−154.9 to 40.2)	R	86.5 (35.6 to 97.2)	R	−47.0 (−331.3 to 49.9)	R	−38.2 (−40.1 to −36.4)	R
27	26.1 (19.9 to 31.9)	R	53.5 (49.6 to 57.1)	R	8.4 (−65.1 to 49.2)	R	4.3 (−150.6 to 63.4)	R	49.8 (49.8 to 49.8)	R
28	16.0 (0.8 to 28.9)	R	9.0 (−6.8 to 22.5)	R	37.7 (18.9 to 52.1)	R	35.3 (−0.1 to 58.2)	R	n/a	
29	51.2 (44.2 to 57.2)	R	69.6 (65.9 to 72.9)	R	55.8 (17.1 to 76.4)	R	55.3 (−0.2 to 80.0)	R	81.7 (81.6 to 81.8)	R
Mean	43.1		36.5		1.8		24.9		−3.1	

FECR percentages indicating susceptible or suspected of resistance status are set in bold. *Note:* $FECR_1$ and $FECR_2$ used pre- and posttreatment fecal egg counts (FECs) from both treated and untreated animals, but $FECR_1$ used arithmetic means while $FECR_2$ used geometric means. $FECR_3$ was calculated using arithmetic means for pre- and posttreatment FECs from treated animals. $FECR_4$ was calculated using arithmetic means for posttreatment FECs from treated and untreated animals, whereas $FECR_5$ was calculated using FEC estimates from a general linear mixed model

n/a it was not possible to compute the predicted fecal egg counts for this farm as weight information for the control animals was not collected, *R* resistance, *SR* suspected of resistance, *S* susceptible, "−" (in front of a number) indicates that the FEC rose after treatment

For $FECR_5$, the final model of the natural logarithm of posttreatment FECs (Table 5) met the assumptions of both homoscedasticity and normality; however, residual analyses revealed two major outliers. The model was rerun without these observations, but this smaller number of observations did not change the final fit of the model, and there was no reason to omit the observations; hence, the outliers were retained in the final model. The significant variables in the final model included: farm, treatment group, natural logarithm of the pretreatment FECs (Lbefore), and animal weight (as a quadratic function). Also included were the interaction variables Lbefore×farm; Lbefore×treatment; Lbefore×weight; treatment×farm; treatment×weight; and farm×weight. Higher posttreatment FECs were associated with higher pretreatment FECs and fenbendazole treatment (compared with the control group), but lower animal weights, although these main effects were modified by their respective interaction variables. The final model of the natural logarithm of posttreatment FECs explained 77 % of the variation in the outcome when taking into account both fixed and random effects.

Table 3 Fecal egg count reduction (FECR) percentages (with 95 % confidence intervals), and anthelmintic resistance status, following fenbendazole treatment on 20 sheep farms in south-western Ontario, Canada (2010 and 2011), calculated using five different FECR formulae

Farm	$FECR_1$		$FECR_2$		$FECR_3$		$FECR_4$		$FECR_5$	
1	−20.6 (−43.4 to −1.5)	R	−309.3 (−426.1 to −218.5)	R	48.9 (37.2 to 58.3)	R	**96.6 (74.2 to 99.6)**	**SR**	93.4 (93.3 to 93.5)	**SR**
3	23.6 (−92.0 to 69.6)	R	22.6 (−83.8 to 67.4)	R	−79.5 (−256.5 to 9.6)	R	46.2 (−33.6 to 78.3)	R	−3.7 (−4.2 to −3.2)	R
4	81.9 (75.4 to 86.6)	R	84.4 (79.0 to 88.4)	R	84.5 (75.7 to 90.1)	R	78.0 (58.3 to 88.4)	R	82.0 (81.9 to 82.1)	R
8	27.2 (18.8 to 34.8)	R	68.2 (64.7 to 71.2)	R	14.3 (−54.6 to 52.5)	R	0.1 (−115.6 to 53.7)	R	49.5 (49.4 to 49.6)	R
9	12.5 (−11.9 to 31.5)	R	−137.5 (−206.9 to −83.8)	R	−31.8 (−114.4 to 19.0)	R	34.0 (−24.4 to 65.0)	R	n/a	
10	84.2 (78.3 to 88.5)	R	30.2 (−3.9 to 53.1)	R	67.4 (51.6 to 78.0)	R	72.5 (−57.4 to 95.2)	R	16.4 (15.5 to 17.3)	R
11	3.4 (−6.9 to 12.7)	R	−51.9 (−73.3 to −33.2)	R	−1.5 (−71.5 to 40.0)	R	−14.3 (−119.5 to 40.5)	R	41.7 (41.6 to 41.8)	R
12	65.3 (60.2 to 69.8)	R	59.8 (54.4 to 64.6)	R	56.1 (16.1 to 77.0)	R	−22.4 (−192.7 to 48.8)	R	25.4 (25.3 to 25.5)	R
13	81.1 (71.5 to 87.5)	R	**91.1 (86.0 to 94.3)**	**SR**	88.6 (79.5 to 93.7)	R	92.0 (75.1 to 97.4)	R	**91.0 (90.9 to 91.1)**	**SR**
16	73.2 (59.4 to 82.4)	R	66.1 (44.0 to 79.5)	R	51.8 (34.1 to 64.7)	R	90.2 (78.7 to 95.5)	R	86.1 (86.0 to 86.2)	R
18	79.9 (77.5 to 81.9)	R	67.1 (62.7 to 71.0)	R	27.1 (5.0 to 44.0)	R	78.0 (49.7 to 90.4)	R	75.1 (75.0 to 75.2)	R
20	76.4 (72.5 to 79.7)	R	65.4 (61.8 to 68.7)	R	38.7 (25.1 to 49.8)	R	70.8 (−5.6 to 92.0)	R	61.0 (60.9 to 61.1)	R
21	18.5 (−22.7 to 45.8)	R	48.5 (26.9 to 63.8)	R	−10.2 (−548.6 to 81.3)	R	54.7 (−131.2 to 91.1)	R	36.6 (35.7 to 37.5)	R
23	89.7 (87.2 to 91.8)	R	**91.0 (88.9 to 92.8)**	**SR**	76.6 (52.7 to 88.4)	R	79.3 (52.6 to 91.0)	R	83.8 (83.7 to 83.9)	R
24	73.0 (52.5 to 84.7)	R	58.3 (22.7 to 77.6)	R	48.1 (−0.7 to 73.2)	R	80.9 (27.1 to 95.0)	R	75.5 (75.1 to 75.8)	R
25	23.8 (−68.5 to 65.6)	R	37.8 (−22.5 to 68.4)	R	75.8 (47.7 to 88.8)	R	−260.2 (−1579.4 to 22.7)	R	−56.6 (−58.9 to −54.4)	R
26	18.2 (−84.9 to 63.8)	R	−61.9 (−277.6 to 30.6)	R	48.1 (−19.0 to 77.3)	R	−18.1 (−260.4 to 61.3)	R	−1.8 (−3.1 to 40.1)	R
27	44.4 (40.6 to 47.9)	R	47.8 (52.1 to 56.0)	R	31.0 (5.3 to 49.8)	R	27.6 (−88.5 to 72.2)	R	40.0 (39.9 to 40.1)	R
28	66.1 (57.7 to 72.9)	R	71.0 (62.4 to 77.6)	R	74.9 (59.8 to 84.3)	R	60.2 (33.5 − 76.2)	R	n/a	
29	−9.6 (−21.0 to 0.8)	R	−72.7 (−91.1 to −56.1)	R	0.9 (−77.7 to 44.7)	R	23.5 (−63.9 to 64.3)	R	59.3 (59.2 to 59.4)	R
Mean	45.6		13.8		35.5		33.5		47.5	

FECR percentages indicating susceptible or suspected of resistance are set in bold. See Table 2 for description of different FECR methods

n/a it was not possible to compute the predicted fecal egg counts for this farm as weight information for the control animals was not collected, *R* resistance, *SR* suspected of resistance, "−" (in front of a number) indicates that the FEC rose after treatment

Agreement between farm resistance status based on different FECR calculation methods

The Kappa values and categories of agreement (beyond that due to chance alone) between the farm resistance status obtained with the five FECR calculation methods for ivermectin, fenbendazole, and levamisole are shown in Table 6. The Kappa values and categories of agreement varied depending on which FECR calculation methods were compared, and which anthelmintic treatment group was being assessed.

For ivermectin resistance status, agreement among calculation methods (beyond that due to chance) ranged from moderate (e.g., between $FECR_1$ and $FECR_5$, $FECR_2$, and $FECR_3$) to almost perfect (e.g., between $FECR_1$ and $FECR_2$). The overall mean Kappa values for $FECR_3$ and $FECR_4$ (the logistically easier methods and simpler formulae) when comparing them with the other FECR methods, were 0.52 (moderate) and 0.66 (substantial), respectively, whereas the mean Kappa values when comparing them with only $FECR_1$ and $FECR_2$ were 0.48 (moderate) and 0.66 (substantial), respectively.

Kappa values for fenbendazole treatment could not be computed for $FECR_1$ and $FECR_3$, as there was no variation

in the classification (i.e., all farms were classified as fenbendazole resistant using these two FECR calculation methods).

For levamisole resistance status, agreement among calculation methods (beyond that due to chance alone) ranged from fair (e.g., between $FECR_2$ and $FECR_3$ and $FECR_2$ and $FECR_4$) to almost perfect (e.g., between $FECR_3$ and $FECR_5$). The overall mean Kappa value for both $FECR_3$ and $FECR_4$, when comparing them with all other FECR methods, was 0.74 (substantial), whereas the mean Kappa value when comparing them with only $FECR_1$ and $FECR_2$ was 0.54 (moderate).

When $FECR_3$ and $FECR_4$ were re-computed without a BCT, the farm classifications for resistance changed somewhat for levamisole, leading to different Kappa values (see Table 7), whereas the AR classifications for ivermectin and fenbendazole did not change, leading to identical Kappa values as those in Table 6. Farms 25 and 26 were classified as resistant by $FECR_3$ and $FECR_4$ when a BCT was added, in agreement with $FECR_1$, $FECR_2$, and $FECR_5$ (Table 4), but they were classified as susceptible when no BCT was added. Consequently, the Kappa values between all pair-wise combinations including either $FECR_3$ or $FECR_4$ decreased with no

Table 4 Fecal egg count reduction (FECR) percentages (with 95 % confidence intervals) and anthelmintic resistance status, following levamisole treatment on 17 sheep farms in south-western Ontario, Canada (2010 and 2011), calculated using five different FECR formulae

Farm	$FECR_1$		$FECR_2$		$FECR_3$		$FECR_4$		$FECR_5$	
1	**94.8 (82.6 to 98.4)**	**R**	−21.0 (−301.4 to 63.5)	**R**	97.8 (91.9 to 99.4)	S	99.3 (97.0 to 99.8)	S	95.2 (95.1 to 95.3)	S
4	98.4 (94.6 to 99.5)	S	98.0 (93.3 to 99.4)	S	98.6 (97.9 to 99.1)	S	98.2 (97.3 to 98.9)	S	97.2 (97.1 to 97.3)	S
8	99.8 (99.2 to 99.9)	S	99.6 (98.6 to 99.9)	S	99.7 (99.5 to 99.8)	S	99.7 (99.8 to 99.5)	S	99.4 (99.3 to 99.5)	S
9	97.5 (92.9 to 99.1)	S	92.5 (77.5 to 97.5)	R	96.3 (93.9 to 97.7)	S	99.2 (98.5 to 99.5)	S	n/a	
10	**93.8 (86.2 to 97.2)**	**R**	**75.8 (42.2 to 89.9)**	**R**	**87.1 (76.7 to 92.9)**	**R**	**92.0 (57.9 to 98.5)**	**R**	**59.0 (58.7 to 59.4)**	**R**
11	99.4 (98.2 to 99.8)	S	99.1 (97.3 to 99.7)	S	99.4 (98.9 to 99.7)	S	99.5 (99.3 to 99.7)	S	99.3 (99.2 to 99.4)	S
12	99.7 (99.1 to 99.9)	S	99.3 (97.6 to 99.8)	S	99.6 (99.4 to 99.8)	S	99.3 (98.7 to 99.6)	S	98.5 (98.4 to 98.6)	S
13	97.2 (94.1 to 98.7)	S	**94.8 (88.8 to 97.6)**	**R**	98.3 (97.7 to 98.8)	S	96.9 (92.1 to 98.8)	S	95.2 (95.1 to 95.3)	S
16	98.5 (95.0 to 99.6)	S	**94.9 (82.8 to 98.5)**	**R**	97.3 (94.6 to 98.7)	S	97.8 (95.7 to 98.9)	S	95.8 (95.7 to 95.9)	S
18	99.9 (99.74 to 99.97)	S	99.7 (99.2 to 99.9)	S	99.7 (99.4 to 99.8)	S	99.7 (99.4 to 99.8)	S	99.0 (98.9 to 99.1)	S
20	99.8 (99.3 to 99.9)	S	99.1 (96.9 to 99.7)	S	99.5 (99.0 to 997)	S	99.5 (98.5 to 99.8)	S	98.3 (98.2 to 98.4)	S
23	99.7 (99.0 to 99.9)	S	99.1 (97.4 to 99.7)	S	99.2 (97.9 to 99.7)	S	99.4 (98.9 to 99.6)	S	98.2 (98.1 to 98.3)	S
25	**75.5 (−3.6 to 94.2)**	**R**	**9.4 (−262.4 to 77.3)**	**R**	**92.2 (56.1 to 98.6)**	**R**	**68.1 (20.6 to 87.2)**	**R**	**31.4 (30.6 to 32.4)**	**R**
26	**88.9 (54.3 to 97.3)**	**R**	**55.7 (−76.6 to 88.9)**	**R**	**92.9 (79.9 to 97.5)**	**R**	**77.1 (46.0 to 90.3)**	**R**	**60.5 (60.1 to 61.0)**	**R**
27	98.4 (97.5 to 99.0)	S	99.3 (98.9 to 99.6)	S	**98.0 (83.2 to 99.8)**	**SR**	96.5 (78.2 to 99.4)	SR	98.9 (98.8 to 99.0)	S
28	**91.6 (87.3 to 94.5)**	**R**	98.3 (97.3 to 98.9)	S	**93.8 (66.3 to 98.9)**	**R**	**92.4 (56.1 to 98.7)**	**R**	n/a	
29	99.3 (97.6 to 99.8)	S	97.7 (92.4 to 99.3)	S	99.3 (98.7 to 99.7)	S	99.4 (99.1 to 99.6)	S	99.2 (99.1 to 99.3)	S
Mean	96.0		81.8		97.0		94.9		88.0	

FECR percentages indicating resistance status are set in bold. See Table 2 for description of different FECR methods

n/a it was not possible to compute the predicted fecal egg counts for this farm as weight information for the control animals was not collected, *R* resistance, *SR* suspected of resistance, *S* susceptible, "−" (in front of a number) indicates that the FEC rose after treatment

BCT. The overall mean Kappa values of $FECR_3$ and $FECR_4$, when comparing them with all other FECR methods, were 0.41 (moderate) and 0.31 (fair), respectively. The mean Kappa values when comparing them with only $FECR_1$ and $FECR_2$ were 0.21 (fair) and 0.11 (slight), respectively (Table 7). Overall, $FECR_4$ showed the highest agreement with all the other FECR methods, particularly when levels of resistance were low and a BCT was used

Discussion

The results of this study lend some support to the view that different methods of calculating FECR (based on different formulae, and/or based on whether a BCT was added or not) may lead to different classifications of a farm's AR status, at least where low levels of resistance are present. Despite the different classifications produced by the different calculation methods, the $FECR_4$ method appeared to have empirical advantages within this study population, adding to the conceptual and practical reasons for small ruminant clinicians to consider using it when conducting FECRTs on farms, as explained below. Moreover, this manuscript is novel in that it analyzes typical commercial flock data in a temperate continental climate; therefore this evidence may be particularly relevant to clients of small ruminant clinicians in similar climates. We also describe a formula to calculate the 95 % CI for $FECR_1$, $FECR_2$, and $FECR_3$, using asymptotic approximation, that takes into account the correlation between pre- and posttreatment fecal samples collected from the same animals.

In the work described here, for both the ivermectin and fenbendazole treatment groups, almost all farms (93 to 97 % and 88 to 100 % of the farms, respectively) were classified as having AR, regardless of the FECR calculation used (Tables 2 and 3). Moreover, the addition of a BCT to $FECR_3$ and $FECR_4$ did not influence the farm classification resistance status for these two calculation methods. By contrast, the levamisole FECR percentages obtained with the different calculation methods showed greater heterogeneity, compared with the other two anthelmintic treatment groups. The reason for this heterogeneity in FECR percentages could partly be explained by the overall low levels of levamisole resistance present in the studied sheep flocks, which led to a clustering of results and consequent right skewness of the posttreatment data, as suggested by Levecke et al. (2012a). Greater heterogeneity of FECR results for levamisole was observed for farms 25 and 26 specifically, where ewe lambs in their first grazing season were used for the FECRT; young lambs could not be used as they were kept indoors during their first summer because of predator concerns. Consequently, many of these ewe lambs were heavier (50–80 kg) than the lambs on

Table 5 General linear mixed model for the natural logarithm of the posttreatment fecal egg counts (eggs per gram) for 29 sheep farms in south-western Ontario, Canada (2010 and 2011)

	Estimate (95 % CI)	F value	Z value	P value
Fixed effects				
Intercept	8.069 (5.989 to 10.148)	7.61		<0.001
Lbefore	0.298 (0.113 to 0.482)	7.39		0.007
Weight	−0.165 (−0.232 to −0.097)	17.18		<0.001
Weight×weight	0.001 (0.0004 to 0.0015)	11.32		0.001
Treatment		3.77		0.010
Control	Reference			
Ivermectin	−0.244 (−1.887 to 1.399)			
Fenbendazole	2.750 (0.956 to 4.544)			
Levamisole	−0.148 (−1.672 to 1.377)			
Farm	−[a] (−[a])	4.90		<0.001
Lbefore×weight	0.005 (0.001 to 0.010)	5.60		0.002
Lbefore×treatment	−[b] (−[b])	58.52		<0.001
Lbefore×farm	−[b] (−[b])	4.40		<0.001
Weight×treatment	−[b] (−[b])	8.50		<0.001
Weight×farm	−[b] (−[b])	1.79		0.007
Farm×treatment	−[b] (−[b])	4.23		<0.001
Random effects				
Treatment−control	1.459 (1.246 to 1.733)		11.90	<0.001
Treatment−ivermectin	1.444 (1.240 to 1.705)		12.32	<0.001
Treatment−fenbendazole	1.453 (1.222 to 1.757)		10.81	<0.001
Treatment−levamisole	0.187 (0.151 to 0.238)		8.59	<0.001

CI confidence interval, *Lbefore* natural logarithm of the pretreatment fecal egg counts (eggs per gram)

[a] These coefficients are not reported in the table as a separate coefficient was provided for each of the 29 farms included in the mode.

[b] When interpreting these variables, there is not just one coefficient to consider because these variables are involved in interactions and are categorical. The total effect for each variable is the combination of the relevant coefficients for the main effects and the interacting categories

other farms in the study (10–45 kg). It has been hypothesized that heavier animals may carry more resistant parasites because they are more likely to be underdosed, which in turn may accelerate the development of resistance on farms (Sutherland and Scott 2010). However, all animals in our study were weighed individually prior to treatment to remove the risk of underdosing. Furthermore, the FECR$_5$ method took into account the animals' weights in the model (weight included as a confounder), which influenced the overall estimate provided by the GLMM. As indicated in Table 5, weight was negatively correlated with the FECR percentage; lower FECR percentages were estimated for the heavier animals.

The addition of a BCT when using either FECR$_3$ or FECR$_4$ influenced the classification of two farms' levamisole resistance status (farms 25 and 26), and the overall agreement between the FECR methods improved when a BCT was added. A BCT of 25 was chosen as the McMaster technique used had a minimum detection limit of 50 epg. More sensitive diagnostic tests have been described in the literature, such as the FECPAK (detection limit=10 epg) or the FLOTAC technique (detection limit=1–2 epg) (El-Abdellati et al. 2010). However, these techniques are either more expensive, more labor intensive, and/or require additional training or equipment, which has so far limited their uptake and application. Moreover, recent work by Levecke et al. (2012b) showed that when the FECs are low and the drug efficacy is high, the

anthelmintic efficacy is likely to be underestimated, regardless of the sensitivity of the test.

Overall, the FECR results obtained for the three different treatment groups are in agreement with Miller et al. (2006), who suggested that the FECRT is effective at diagnosing resistance when AR is present at high levels but is more variable when the drug efficacy ranges around 90–95 % and/ or when low levels of resistance are present, as was the case for levamisole observed here. Higher variability was particularly evident between FECR calculation methods for levamisole resistance when many FECs of 0 epg were present in the pretreatment FECs, likely leading to underestimates of overall treatment efficacy (Dobson et al. 2009).

Use of the simpler formulae (i.e., inclusion of pre- and posttreatment data from treated animals only [FECR3], or inclusion of only posttreatment data from treated and control animals [FECR$_4$]) seemed not to be very influential on the FECR percentages, regardless of the level of resistance, as demonstrated by their moderate to substantial Kappa values with FECR$_1$ and FECR$_2$, respectively. These results are in agreement with a study by McKenna (2006), which reported that different FECR formulae (using only arithmetic means) detected a similar number of cases of AR, and therefore suggested that the simpler formulae could be a suitable alternative to the more complex formulae which require dual samples from both treated and untreated animals.

Table 6 Weighted Kappa values (and standard error) for the agreement between the farm resistance statuses, based on five different methods for calculating fecal egg count reduction (FECR) percentages following (a) ivermectin, (b) fenbendazole, and (c) levamisole treatment for 29, 20, and 17 sheep farms, respectively, in south-western Ontario, Canada (2010 and 2011)

	$FECR_1$	$FECR_2$	$FECR_3$	$FECR_4$
Ivermectin				
$FECR_2$	1.00 (0.19)	–	–	–
	Almost perfect			
$FECR_3$	0.48 (0.13)	0.48 (0.13)	–	–
	Moderate	Moderate		
$FECR_4$	0.66 (0.12)	0.66 (0.12)	0.65 (0.18)	–
	Substantial	Substantial	Substantial	
$FECR_5$	0.48 (0.13)	0.48 (0.13)	0.46 (0.19)	0.65 (0.18)
	Moderate	Moderate	Moderate	Substantial
Fenbendazole				
$FECR_2$	n/a	–	–	–
$FECR_3$	n/a	n/a	–	–
$FECR_4$	n/a	−0.07 (0.21)	n/a	–
		Poor		
$FECR_5$	n/a	0.44 (0.24)	n/a	0.64 (0.22)
		Moderate		Substantial
Levamisole				
$FECR_2$	0.49 (0.23)	–	–	–
	Moderate			
$FECR_3$	0.78 (0.23)	0.29 (0.22)	–	–
	Substantial	Fair		
$FECR_4$	0.78 (0.23)	0.29 (0.22)	1.00 (0.22)	–
	Substantial	Fair	Perfect	
$FECR_5$	0.81 (0.25)	0.55 (0.23)	0.90 (0.24)	0.90 (0.24)
	Almost perfect	Moderate	Almost perfect	Almost perfect

n/a it was not possible to compute the Kappa agreement as there was no variation in the farm classification of fenbendazole resistance status using $FECR_1$ and $FECR_3$ (i.e., all farms were classified as fenbendazole resistant)

Table 7 Weighted Kappa values (and standard error) for the agreement between the farm resistance statuses, based on five different methods for calculating fecal egg count reduction (FECR) percentages following levamisole treatment on 17 sheep farms, respectively, in south-western Ontario, Canada (2010 and 2011), with $FECR_3$ and $FECR_4$ calculations not including a bias correction term

	$FECR_1$	$FECR_2$	$FECR_3$	$FECR_4$
Levamisole				
$FECR_2$	Unchanged	–	–	–
	From Table 6			
$FECR_3$	0.42 (0.21)	−0.001 (0.18)	–	–
	Moderate	Poor		
$FECR_4$	0.31 (0.18)	−0.09 (0.15)	0.87 (0.20)	–
	Fair	Poor	Almost perfect	
$FECR_5$	Unchanged	Unchanged	0.36 (0.22)	0.17 (0.15)
	From Table 6	From Table 6	Fair	Slight

The timeliness and relevance of our results are indicated by the recent literature published on the topic (Denwood et al. 2010; Levecke et al. 2011; Dobson et al. 2012; McKenna 2013), highlighting the urgency to reach a consensus as to which calculation method should be used. Shortly after the World Association for the Advancement of Veterinary Parasitology (WAAVP)-endorsed $FECR_4$ as the standardized method of calculation (Coles et al. 1992), a second set of WAAVP-endorsed guidelines, recommending the use of geometric means, were published by Wood et al. (1995). The latter guidelines were criticized by McKenna (1997) and Van Wyk and Groeneveld (1997), and as that time there has been an ongoing discussion as to which method is more appropriate (Denwood et al. 2010; Dobson et al. 2012).

Given the dramatic rise and spread of AR reported in recent years (Papadopoulos et al. 2012; Torres-Acosta et al. 2012), regular AR monitoring is fundamental to slow the further development of AR (Abbott et al. 2009). Thus, the FECRT must be as accessible and straightforward as possible to facilitate its inclusion within routine flock health checks and to improve surveillance for the development of AR. From a practical point of view, and based on results from this study, we therefore recommend the re-introduction of $FECR_4$ as the standardized method used by clinicians for several empirical, practical, and conceptual reasons. Firstly, this method showed the best overall agreement with other FECR calculations when a BCT was added (substantial agreement for both ivermectin and levamisole). Since this method does not require the inclusion of pretreatment FECs, it is less laborious (El-Abdellati et al. 2010) and more acceptable to sheep producers (McKenna 2006), facilitating its implementation by small ruminant clinicians. Moreover, inclusion of the control group makes $FECR_4$ conceptually less likely to be prone to selection bias, for the same reasons we include a control group in a randomized controlled trial. The random allocation to treatment and control groups usually leads to very similar groups of animals, except that one group is getting a treatment, thus minimizing selection bias, information bias, and confounding bias (Dohoo et al. 2009). Conversely, inclusion of pretreatment FECs in the $FECR_3$ calculation method would only minimize information bias and confounding bias (i.e., the animals are acting as their own controls, with likely similar measurement errors and confounders before and after testing), but not minimize selection bias.

From a research point of view, however, $FECR_1$ and $FECR_2$ are better options since they combine all the following advantages: (i) the use of both treated and control groups; (ii) random allocation of animals to these groups; and (iii) the use of baseline and posttreatment FECs. Together, these FECR

components allow researchers to see the temporal changes in FECs on farms in untreated and treated animals, making it possible to confirm that any changes observed in the treated animals are in fact due to the treatment, as recommended by Torgerson et al. (2005).

Different CIs are described for $FECR_1$ and $FECR_2$, compared with $FECR_3$, as the former calculation methods require an odds ratio approach while the latter method requires a straight ratio approach (Kahn and Sempos 1989). Nonetheless, both approaches are based on asymptotic approximation and adjust for the correlation between pre- and posttreatment FECs from the same animal (Mood et al. 1974). We were unable to perform simulation studies to verify the coverage of these CIs because more information from a larger number farms would be required to inform parameters on sheep-to-sheep and farm-to-farm variances, correlation structures between pre- and posttreatment FECs, and range of drug efficacies. Such simulation studies could be an area for future research.

Comparisons between $FECR_1$ and $FECR_2$ produced Kappa values of 1.00 for ivermectin and 0.49 for levamisole, suggesting that they produced identical classifications for the high levels of resistance to ivermectin, but moderate agreement in classifications for the low levels of resistance to levamisole. Therefore, we cannot say the two methods produced identical results all the time. However, we cannot say which is better because we do not have some gold standard measurement of GIN parasitism upon which to evaluate them. Future research with a gold standard is recommended to clarify whether one of these two calculation methods is preferred.

In conclusion, the different FECR calculation methods evaluated in this study did not provide consistent FECR percentages following treatment with ivermectin, fenbendazole, or levamisole. The agreement in farm AR classifications (beyond that due to chance alone) between the methods was largely influenced by: (i) the level of AR and (ii) whether a BCT was used in the FECR formulae—use of a BCT improved overall agreement between the FECR methods, especially when low levels of resistance were present (e.g., the levamisole group in our study). By contrast, use of the simpler formulae (i.e., only posttreatment data from treated and control animals or pre- and posttreatment data from treated animals only) versus use of both pre- and posttreatment data from both treated and untreated animals, was not very influential on the agreement in farm AR classifications in this study population. Overall, $FECR_4$ showed the highest agreement with all the other FECR methods, particularly when levels of resistance were low and a BCT was used. We therefore recommend that small ruminant clinicians in temperate continental climates use the $FECR_4$ formula described by Coles et al. (1992), which uses posttreatment FECs from treated and control animals, when AR levels are suspected to be quite high (>85 %) or quite low (<35 %), the range of FECR percentages found in our study. This recommendation would minimize the cost and labor associated with the FECRT, while allowing for standardized comparisons to be made between different farms and studies, with minimal risk for selection bias, information bias and confounding bias. For researchers, we recommend the use of $FECR_1$ or $FECR_2$, because the inclusion of both pre- and posttreatment FECs, and use of randomly allocated animals in treatment and control groups, makes them mathematically more likely to estimate the true anthelmintic efficacy.

References

Abbott KA, Taylor MA, Stubbings LA (2009) Sustainable worm control strategies for sheep. 3rd edn. A technical manual for veterinary surgeons and advisers. 51 pp. Sustainable Control of Parasites in Sheep (SCOPS) and printed by: Context Publications

Cabaret J, Berrag B (2004) Fecal egg count reduction test for assessing anthelmintic efficacy: average versus individually based estimations. Vet Parasitol 121:105–113

Coles GC, Bauer C, Borgsteede FHM, Geerts S, Klei TR, Taylor MA, Waller PJ (1992) World Association for the Advancement of Veterinary Parasitology (W.A.A.V.P.) methods for the detection of anthelmintic resistance in nematodes of veterinary importance. Vet Parasitol 44:35–44

Coles GC, Jackson F, Pomroy WE, Prichard RK, von Samson-Himmelstjerna G, Silvestre A, Taylor MA, Vercruysse J (2006) The detection of anthelmintic resistance in nematodes of veterinary importance. Vet Parasitol 136:167–185

Dash K, Hall K, Barger IA (1988) The role of arithmetic and geometric worm egg counts in faecal egg count reduction tests and in monitoring strategic drenching programs in sheep. Aust Vet J 65:66–68

Denwood M, Reid SWJ, Love S, Nielsen MK, Matthews L, McKendrick IJ, Innocent GT (2010) Comparison of three alternative methods for analysis of equine faecal egg count reduction test data. Prev Vet Med 93:316–323

Dobson RJ, Sangster NC, Besier RB, Woodgate RG (2009) Geometric means provide a biased efficacy result when conducting a faecal egg count reduction test (FECRT). Vet Parasitol 161:162–167

Dobson RJ, Hosking BC, Jacobson CL, Cotter JL, Besier RB, Stein PA, Reid SA (2012) Preserving new anthelmintics: a simple method for estimating faecal egg count reduction test (FECRT) confidence limits when efficacy and/or nematode aggregation is high. Vet Parasitol 186:79–92

Dohoo I, Martin W, Stryhn H (2009) Veterinary Epidemiologic Research-Model-Building Strategies (ISBN: 978-0-919013-60-5). 2nd edn. VER Inc., Charlottetown, Prince Edward Island, Canada, C1A 8X5. pp. 365–390

El-Abdellati A, Charlier J, Geldhor P, Levecke B, Demeler J, von Samson-Himmelstjerna G, Claerebout E, Vercruysse J (2010) The use of a simplified faecal egg count reduction test for assessing

anthelmintic efficacy on Belgian and German cattle farms. Vet Parasitol 169:352–357

Falzon LC, Menzies PI, Shakya KP, Jones-Bitton A, Vanleeuwen J, Avula J, Stewart H, Jansen JT, Taylor MA, Learmount J, Peregrine AS (2013) Anthelmintic resistance in sheep flocks in Ontario, Canada. Vet Parasitol 193:150–162

Fulford AJC (1994) Dispersion and bias: can we trust geometric means? Parasitol Today 10:446–448

Jackson F, Coop RL (2000) The development of anthelmintic resistance in sheep nematodes. Parasitology 120:S95–S107

Kahn HA, Sempos CT (1989) Statistical methods in epidemiology. Oxford Univ. Press, New York, 292 pp

Kaplan RM (2004) Drug resistance in nematodes of veterinary importance: a status report. Trends Parasitol 20:477–481

Kaplan RM, Vidyashankar AN (2012) An inconvenient truth: global worming and anthelmintic resistance. Vet Parasitol 186:70–78

Knox MR, Besier RB, Le Jambre LF, Kaplan RM, Torres-Acosta JFJ, Miller J, Sutherland I (2012) Novel approaches for the control of helminth parasites of livestock VI: summary of discussion and conclusion. Vet Parasitol 186:143–149

Levecke B, Rinaldi L, Charlier J, Maurelli MP, Morgoglione ME, Vercruysse J, Cringoli G (2011) Monitoring drug efficacy against gastrointestinal nematodes when faecal egg counts are low: do the analytic sensitivity and the formula matter? Parasitol Res 109:953–957

Levecke B, Dobson RJ, Speybroeck N, Vercruysse J, Charlier J (2012a) Novel insights in the fecal egg count reduction test for monitoring drug efficacy against gastrointestinal nematodes of veterinary importance. Vet Parasitol 188:391–396

Levecke B, Rinaldi L, Charlier J, Maurelli MP, Bosco A, Vercruysse J, Cringoli G (2012b) The bias, accuracy and precision of faecal egg count reduction test results in cattle using McMaster, Cornell-Wisconsin and FLOTAC egg counting methods. Vet Parasitol 188:194–199

Littell RC, Milliken GA, Stroup WW, Wolfinger RD (1996) SAS® system for mixed model. SAS Institute Inc, Cary, 633 pp

McKenna PB (1997) Use of arithmetic and geometric means in the calculation of anthelmintic efficacy. Vet Record 141:472–473

McKenna PB (2006) A comparison of faecal egg count reduction test procedures. N Z Vet J 54:202–203

McKenna PB (2013) Are multiple pre-treatment groups necessary or unwarranted in faecal egg count reduction tests in sheep? Vet Parasitol 196:433–437

Mejia MD, Fernandez IB, Schmidt EE, Cabaret J (2003) Multispecies and multiple anthelmintic resistance on cattle nematodes in a farm in Argentina: the beginning of high resistance? Vet Res 34:461–467

Miller CM, Waghorn TS, Leathwick DM, Gilmour ML (2006) How repeatable is a faecal egg count reduction test? N Z Vet J 54:323–328

Ministry of Agriculture, Fisheries and Food (1986) Manual of veterinary parasitological laboratory techniques. Technical Bulletin No. 18. HMSO, London, 124 pp

Mood AM, Graybill FA, Boes DC (1974) Introduction to the theory of statistics, 3rd edn. McGraw-Hill, New York, p 480

Papadopoulos E, Gallidis E, Ptochos S (2012) Anthelmintic resistance in sheep in Europe: a selected review. Vet Parasitol 189:85–88

Presidente PJA (1985) Methods for detection of resistance to anthelmintics. In: Anderson N, Waller PJ (eds) Resistance in nematodes to anthelmintic drugs. CSIRO Division of Animal Health, Glebe, NSW, Australia, pp 13–28

Prichard RK, Hall CA, Kelly JD, Martin ICA, Donald AD (1980) The problem of anthelmintic resistance in nematodes. Aust Vet J 56:239–250

Sargison ND (2008) Sheep flock health: a planned approach. Blackwell Publishing Ltd, Oxford, UK, pp 149–191

Smothers CD, Sun F, Dayton AD (1999) Comparison of arithmetic and geometric means as measures of a central tendency in cattle nematode populations. Vet Parasitol 81:211–224

Sutherland I, Scott I (2010) Gastrointestinal nematodes of sheep and cattle. Biology and control. John Wiley and Sons Ltd, West Sussex, UK, pp 1–60

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

Torres-Acosta JFJ, Mendoza-de-Gives P, Aguilar-Caballero AJ, Cuéllar-Ordaz JA (2012) Anthelmintic resistance in sheep farms: update of the situation in the American continent. Vet Parasitol 189:89–96

Van Wyk JA, Groeneveld HT (1997) Comments on the paper World Association for the Advancement of Veterinary Parasitology (W.A.A.V.P.) second edition of guidelines for evaluating the efficacy of anthelmintics in ruminants (bovine, ovine, caprine). Vet Parasitol 70:283–288

Wood IB, Amaral NK, Bairden K, Duncan JL, Kassai T, Malone JB Jr, Pankavich JA, Reinecke RK, Slocombe O, Taylor SM, Vercruysse J (1995) World Association for the Advancement of Veterinary Parasitology (W.A.A.V.P.) second edition of guidelines for evaluating the efficacy of anthelmintics in ruminants (bovine, ovine, caprine). Vet Parasitol 58:181–213