Veterinary Parasitology 205 (2014) 385–388

Contents lists available at ScienceDirect

Veterinary Parasitology

journal homepage: www.elsevier.com/locate/vetpar

Short communication

# Proficiency testing assessments for nematode worm egg counting based on Poisson variation

A.J. van Burgel[a,*], J. Lyon[a], R.B. Besier[a], D.G. Palmer[b]

[a] Department of Agriculture and Food, 444 Albany Highway, Albany 6330, WA, Australia
[b] Animal Health Laboratories, Department of Agriculture and Food, 3 Baron Hay Court, South Perth 6151, WA, Australia

ARTICLE INFO

*Article history:*
Received 4 February 2014
Received in revised form 19 June 2014
Accepted 29 June 2014

*Keywords:*
Proficiency testing
Poisson distribution
Quality assurance
Laboratory
Faecal egg count
Worm egg count

ABSTRACT

The Poisson distribution provides an appropriate model for the variation within laboratories in worm egg counting. This is demonstrated by the results from annual quality assessment exercises in which laboratories in Australia tested multiple samples from the same mixtures prepared using different worm egg densities. Confidence intervals based on simulations using the Poisson distribution are recommended in the analysis of the results to identify laboratories showing significant bias or overdispersion, hence indicating possible procedural errors.



## 1. Introduction

There is an increasing requirement for veterinary diagnostic and analytical laboratories to participate in proficiency testing schemes for quality control and accreditation purposes. For many quantitative laboratory measurements the assessment of inter-laboratory programs is well established. It is usually based on robust statistics and the use of *z*-scores calculated from the variation between laboratories to assess participants' performances. Outliers are defined as laboratories with *z*-scores exceeding three in magnitude (Thompson et al., 2006). There is no established assessment proficiency testing method for the measurement of worm egg counts in faeces. Sustainable worm control practices rely heavily on monitoring worm egg counts as management tools (Besier, 2004; Karlsson and Greeff, 2006).

The Department of Agriculture and Food Western Australia has been offering a worm egg count quality assurance (QA) scheme for private and government veterinary parasitology laboratories in Australia since 1996 to ensure that the sheep industry has access to a high quality worm egg count service. Until 2005, the assessment was based on robust statistics and *z*-scores. A re-examination of the data found that *z*-score values greater than three in magnitude were extremely rare even when obvious outliers or calculation errors occurred to the extent that the affected laboratory result lacked the fitness-for-purpose criteria (Thompson et al., 2006). We here propose a new assessment method based on Poisson variation, with concerns raised where laboratories are identified as showing significant bias or variation conflicting with Poisson variation. The method is in line with the comment about quality control in a recent paper that very eloquently describes the appropriateness of Poisson variation for the

* Corresponding author. Tel.: +61 8 9892 8550; fax: +61 8 98412707.
*E-mail address:* andrew.vanburgel@agric.wa.gov.au (A.J. van Burgel).

http://dx.doi.org/10.1016/j.vetpar.2014.06.038

distribution of replicate sample egg counts (Torgerson et al., 2012).

## 2. Materials and methods

### 2.1. Sample preparation

Faecal samples were spiked with a known number of eggs to assure homogeneity of samples sent to participating laboratories (Thompson et al., 2006). Faecal material was collected from worm free donor sheep, and stored at −15 °C until required. Sufficient 2 g lots (±0.15 g) were weighed into 70 mL containers to allocate 13 samples to each participating laboratory.

Nematode eggs were purified from faeces of sheep infected with *Haemonchus contortus* and the egg concentration determined. Each participating laboratory received a set of 13 test samples, 5 samples were "spiked" with 600 eggs to achieve 300 eggs per gram (EPG) in the test samples and 5 spiked with 2000 eggs (equivalent to 1000 EPG). Three samples were left un-spiked. Samples in each group (1000 EPG, 300 EPG, 0 EPG) were numbered and randomly allocated to laboratories. Ten samples from each group were randomly selected and counted twice to ensure that homogeneity of samples was within 10% of the mean value.

Samples were sent to laboratories by overnight courier with cooling blocks so that the samples remained cool during transport. Laboratory operators were provided with instructions on how to process the samples. More than one operator could participate provided that each operator filled and read their own counting chambers. A modified McMaster technique (Hutchinson, 2008) was recommended, and a copy of this was sent to each laboratory, but laboratories were encouraged to conduct the assessment with the method they currently use.

A total of 83 laboratories from all over Australia participated during the period 2002 to 2012, but the number in any one year varied from 41 to 55. There were up to eight operators per laboratory and the total number of operators varied from 72 to 93 per year. The laboratories were accredited government laboratories, commercial veterinary diagnostic laboratories, veterinary practices or private worm egg count providers. Participation was voluntary and based on self-assessment principles.

### 2.2. Statistical evaluation

The Poisson distribution is used to model the number of eggs counted, prior to multiplication by the conversion factor to obtain the worm egg count. The Poisson distribution is characterized by equality between its mean and variance. Therefore the ratio of the variance divided by the mean, or index of dispersion (ID) is equal to one. When the variance is greater than or less than the mean this is referred to as 'overdispersion' and 'underdispersion', respectively. Given the observed ID is on some occasions many times greater than one, the square root of ID is used for presentation purposes.

For each operator in each exercise the square root ID was calculated for both the 300 EPG and the 1000 EPG sets of results (5 counts each). The distribution of the square root IDs for each laboratory across all operators and exercises was compared using the Kolmogorov Smirnov test in GenStat (Payne, 2009) to the distribution of square root IDs assuming a Poisson distribution. The later was obtained by calculating the square root ID for 50,000 simulations of sets of 5 egg counts, with each set from a Poisson distribution with mean obtained from a uniform distribution with minimum 3 and maximum 16. Given a conversion factor of 50, this corresponds to a mean worm egg count of between 150 EPG and 800 EPG. These were chosen because the average efficiency of the 300 samples was above 50% in all exercises and the average efficiency of the 1000 EPG samples was below 80% in all but one exercise. However it was also observed by simulation that the distribution of square root ID hardly changes with the mean. For example, the 99% upper limit on the square root ID is 1.80 and 1.82 for a Poisson distribution with mean 3 and 16, respectively. This feature is helpful because it allows repeated worm egg counts to be compared to Poisson variation without needing to know the true mean.

The new assessment method was implemented from 2006 and applied retrospectively to the annual QA rounds from 2002 to 2005. The method uses simulated 95% and 99.73% confidence intervals (the later corresponding to a $z$-score of 3) for the average EPG and square root ID of the five 300 EPG and five 1000 EPG results. The confidence intervals were calculated after one million simulations of the QA worm egg counting process using the statistical software package R (R Development Core Team, 2009). Each simulation involved generating random values from the Poisson distribution to obtain the number of eggs counted for each of the 300 EPG and 1000 EPG samples, from which the various summary statistics were calculated. The desired confidence intervals for each summary statistic were then obtained by the relevant percentiles (i.e. 2.5% and 97.5% for a 95% confidence interval) of the distribution of the one million simulated values of the summary statistic. For example the square root ID of five counts with Poisson variation has an upper 99.73% confidence limit of 2.0.

Given a Poisson distribution with known true mean, exact confidence intervals for the sample mean can be easily calculated. However, while the initial samples were prepared at 300 EPG and 1000 EPG, experience indicates that not all eggs are detected with the technique so that the true mean obtained by correctly applying the worm egg count is not known. Simulation was therefore used with the true mean treated as a random variable with mean equal to the consensus mean of all laboratories and standard deviation equal to the standard error of the consensus mean.

## 3. Results

The distribution of the observed square root IDs for both the 300 EPG and 1000 EPG samples is compared in Fig. 1(A) to that expected from an underlying Poisson distribution. It shows that while most of the square root IDs are consistent with what would be expected from a Poisson distribution, there is some evidence of overdispersion with 5% of the square root IDs above the 99% quantile of the simulated distribution.