Veterinary Parasitology 247 (2017) 85–92

Contents lists available at ScienceDirect

Veterinary Parasitology

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

Research paper

Evaluation of accuracy and precision of a smartphone based automated parasite egg counting system in comparison to the McMaster and Mini-FLOTAC methods

J.A. Scare[a,*], P. Slusarewicz[a,b], M.L. Noel[a], K.M. Wielgus[c], M.K Nielsen[a]

[a] M.H. Gluck Equine Research Center, Department of Veterinary Science, University of Kentucky, Lexington, KY, USA
[b] MEP Equine Solutions, 3905 English Oak Circle, Lexington, KY 40514, USA
[c] College of Veterinary Medicine, Lincoln Memorial University, Harrogate, TN, USA

ARTICLE INFO

Keywords:
Egg count
Smartphone
Automated
Strongyle
McMaster
Mini-FLOTAC

ABSTRACT

Fecal egg counts are emphasized for guiding equine helminth parasite control regimens due to the rise of anthelmintic resistance. This, however, poses further challenges, since egg counting results are prone to issues such as operator dependency, method variability, equipment requirements, and time commitment. The use of image analysis software for performing fecal egg counts is promoted in recent studies to reduce the operator dependency associated with manual counts. In an attempt to remove operator dependency associated with current methods, we developed a diagnostic system that utilizes a smartphone and employs image analysis to generate automated egg counts. The aims of this study were (1) to determine precision of the first smartphone prototype, the modified McMaster and ImageJ; (2) to determine precision, accuracy, sensitivity, and specificity of the second smartphone prototype, the modified McMaster, and Mini-FLOTAC techniques. Repeated counts on fecal samples naturally infected with equine strongyle eggs were performed using each technique to evaluate precision. Triplicate counts on 36 egg count negative samples and 36 samples spiked with strongyle eggs at 5, 50, 500, and 1000 eggs per gram were performed using a second smartphone system prototype, Mini-FLOTAC, and McMaster to determine technique accuracy. Precision across the techniques was evaluated using the coefficient of variation. In regards to the first aim of the study, the McMaster technique performed with significantly less variance than the first smartphone prototype and ImageJ ($p < 0.0001$). The smartphone and ImageJ performed with equal variance. In regards to the second aim of the study, the second smartphone system prototype had significantly better precision than the McMaster ($p < 0.0001$) and Mini-FLOTAC ($p < 0.0001$) methods, and the Mini-FLOTAC was significantly more precise than the McMaster ($p = 0.0228$). Mean accuracies for the Mini-FLOTAC, McMaster, and smartphone system were 64.51%, 21.67%, and 32.53%, respectively. The Mini-FLOTAC was significantly more accurate than the McMaster ($p < 0.0001$) and the smartphone system ($p < 0.0001$), while the smartphone and McMaster counts did not have statistically different accuracies. Overall, the smartphone system compared favorably to manual methods with regards to precision, and reasonably with regards to accuracy. With further refinement, this system could become useful in veterinary practice.

1. Introduction

Strongyles are pervasive gastrointestinal nematodes infecting the horse, where the *Cyathostominae* are the most prevalent (Chapman et al., 2003; Kuzmina, 2012). Routine fecal egg counts (FEC) are recommend to guide regular treatment regimens of equine strongyle parasites by monitoring egg shedding levels and drug efficacy (Nielsen et al., 2016). This necessitates accurate and precise FEC methods.

There are numerous techniques with countless modifications available for performing a FEC. In general, a FEC involves weighing a sample of freshly collected feces and homogenizing it with a flotation medium. The fecal slurry is filtered to remove large debris particles, while the flotation principle separates the eggs from debris for identification and quantification using a microscope. Finally, the number of eggs counted is multiplied by a multiplication factor which is determined by the fecal sample weight, the volume of flotation medium, and the final volume of subsample analyzed. Consequently, the varying detection limits (ranging from 1 to 100 eggs per gram) may cause misinterpretations of FECs, particularly when examining low eggs per gram (EPG) post-treatment samples (Vidyashankar et al., 2012).

* Corresponding author.
E-mail address: Jessica.scare@uky.edu (J.A. Scare).

http://dx.doi.org/10.1016/j.vetpar.2017.10.005
Received 16 April 2017; Received in revised form 1 September 2017; Accepted 11 October 2017

Centrifugal flotation is utilized in some methods, such as the Stoll egg counting technique (Stoll, 1930), the modified Wisconsin sugar flotation technique (Egwang and Slocombe, 1982), and the FLOTAC method (Cringoli et al., 2010). However, the centrifugation step strictly limits the investigator to a laboratory and substantially increases the amount of time required for sample processing (Ballweber et al., 2014). Other techniques, such as the McMaster (MAFF, 1986) and the Mini-FLOTAC (Barda et al., 2013) do not require a centrifuge, but rather rely on passive flotation in specially made slides designed to incorporate flotation chambers. The type of flotation solution and specific gravity used may also contribute to variability. Cringoli et al. (2004) observed the capabilities of 14 different flotation mediums with different specific gravities to recover strongyle eggs in sheep feces. All sucrose-based solutions examined, with a specific gravity between 1.2 and 1.35, floated significantly more eggs than other solutions. The amount of sample volume examined may also cause a source of variability. The two-chamber McMaster slides offer the opportunity examine different sample volumes (0.15, 0.3, 0.5, or 1.0 mL). Cringoli et al. (2004) found the 1.0 mL sample volume to be the most reliable, where the smaller sample volumes over-estimated strongyle EPG in sheep feces. Regardless of the choice of FEC method, the manual counting procedure is operator dependent which represents a substantial source of technical variability (Vidyashankar et al., 2012). One study examined sheep strongyle fecal egg counts performed by farm and laboratory staff using the FECPAK method. Despite having undergone extensive training, the farmers significantly overestimated strongyle egg counts (McCoy et al., 2005). Several recent papers have illustrated the importance of evaluating precision and accuracy when comparing FEC methods (Levecke et al., 2012; Godber et al., 2015; Noel et al., 2017). The importance of egg count precision was emphasized by Carstensen et al. (2013), where the greatest source of variability of FECs was among repeated counts performed on the same subsample.

Several recent studies promote the use of imaging procedures and analysis to identify and quantify parasite eggs, thus reducing or even eliminating the operator dependency of manual counts (Yang et al., 2001; Castañón et al., 2007; Mes et al., 2007; Dogantekin et al., 2008; Ghazali et al., 2013; Suzuki et al., 2013). ImageJ is a publicly available image processing program developed by the National Institute of Health. Its particle analysis tool has proven beneficial for automatically counting cells using image analysis (Grishagin, 2015), and may be applicable for counting parasite eggs. These procedures, however, continue to rely on a microscope as well as a desktop or laptop computer to capture and process the sample images. Most recently, a smartphone-based automated parasite fecal egg counting technique was developed (Slusarewicz et al., 2016). This technology makes use of fluorescent staining and image analysis to identify and count parasite eggs in a sample without the use of a microscope.

The aims of this study were (1) to determine precision of the first smartphone prototype, the modified McMaster, and ImageJ; (2) determine the accuracy, precision, sensitivity, and specificity of the second smartphone prototype, the modified McMaster, and Mini-FLOTAC techniques.

2. Materials and methods

This was a two part study. Part one (2.1) evaluated and compared technique precision of the first smartphone prototype,a modified McMaster, and ImageJ, an image analysis computer software program capable of counting particles in an image. Additionally, the egg counts generated within only the gridded portion of the McMaster slide were compared to egg counts of the entire chamber. Part two (2.2 and 2.3) evaluated and compared technique precision, accuracy, sensitivity, and specificity of the second smartphone prototype, the modified McMaster, and the Mini-FLOTAC. A flow chart layout of the methodology can be found in Fig. 3.

2.1. Precision of computerized egg counting systems and McMaster

Fecal samples from a herd of horses harboring a naturally acquired mixed-species strongyle parasite population were collected for this study section. Quadruplicate egg counts were performed on each prepared sample slurry with each egg counting technique to evaluate precision.

2.1.1. Horses

A total of 22 fecal samples were acquired from a university herd of 22 horses of mixed light breeds housed at the University of Kentucky research farm. The herd consists of 21 mares and 1 stallion, with ages ranging from 4 to 15 years old. Fresh fecal samples were collected into a plastic Ziploc bag, sealed, stored in a refrigerator (4 °C), and analyzed within 3–14 days.

2.1.2. Egg counting methods

The three counting methods used to quantify strongyle eggs in a given image/slide were the smartphone system previously developed by Slusarewicz et al. (2016), ImageJ software (National Institutes of Health (NIH), Bethesda, MD, USA) and manual counting on a McMaster slide.

The Fill-FLOTAC apparatus was used to initially prepare all fecal samples as described by Noel et al. (2017). Briefly, 5 g of feces was weighed out and placed in the Fill-FLOTAC homogenizer and suspended in 45 mL of glucose-NaCl flotation medium with a specific gravity of at least 1.24. A uniform fecal slurry was obtained by homogenizing the flotation medium with the sample. The Fill-FLOTAC was used for its homogenization and filtering capabilities. Next, 5 mL of the filtrate was loaded onto a second filtering mechanism, which consisted of two sequentially stacked cell strainers (pluriSelect Life Science, Leipzig, Germany) of 90 μm and 27 μm pore size. The strainers were fitted onto a 50 mL disposable centrifuge tube using the manufacturer's collar, which contained a vacuum port. A 50 mL syringe was used to draw the filtrate material through the filtering apparatus. Phosphate-buffered saline (PBS) was used to rinse the 27 μm membrane four times to remove any remaining debris. To seal the membrane, positive back pressure was applied with the syringe following the final aspiration and 4 mL of 1% sodium hypochlorite (bleach) was pipetted onto the filter and incubated at room temperature (RT) for two minutes. The bleach was then removed from the filters and the sample was washed with an additional 2 mL of PBS. Subsequently, the 27 μm filter was placed onto a holder-collar covered with parafilm (to provide a vacuum seal). To stain the eggs, 1 mL of 60 μg/ml fluorescein (F)-Chitin Binding Domain (CBD) in 1 x blocking buffer (Vector Laboratories, Burlingame, CA, USA) in PBS was added to the filter, gently agitated, and then allowed to incubate at RT for 2 min. The eggs were then harvested from off the filter with a micropipette using 0.5 mL of flotation medium and placed in 1.5 mL microcentrifuge tubes. Flotation medium was added to each sample to a final volume of 0.6 mL and the egg solution was re-suspended and then transferred into one chamber of a 2-chamber etched McMaster slide (Chalex Corp, Protland, OR). The sample was immediately placed into a PrepOne Sapphire fluorescent gel imager (Embi-Tec, San Diego, CA) to illuminate the fluorescently stained eggs. Lastly, the sample was photographed using an iPhone 5 s fitted with a 7 x macro lens (Olloclip, Huntington Beach, CA, USA) as previously described (Slusarewicz et al., 2016). Four individual McMaster chambers, per fecal slurry, were prepared and analyzed in this way.

Immediately after capturing the image, the same McMaster chamber was used to generate two manual counts using a microscope. First, eggs within the entire McMaster chamber were counted, and secondly only the eggs within the gridded portion were counted.

In order to perform the smartphone and ImageJ counts, the images were exported from the phone to a laptop computer running Windows 8.1 and processed with Photoshop CS6 (Adobe Systems Inc., San Jose, CA, USA). First, images were cropped to obtain two new images. The

first was to include the entire McMaster chamber (2100 × 2100 pixels) and the second to contain the McMaster grid only (1200 × 1200 pixels). Strongyle eggs in each sample image were automatically identified and quantified using a smartphone application, previously described by Slusarewicz et al. (2016). The default settings of the application were used, where RGB range 180–255, particle area 30–130, aspect ratio 1.2-2.4, and circularity 0-0.98.

A third egg count was generated using ImageJ, a computer-based image analysis software program. The color space was set to YUV to remove background noise, while the Y setting was fixed at 200 to highlight the fluorescent eggs. Particle size ranged from 30 to 130, circularity 0-0.98, and an aspect ratio between 1.2 and 2.4. Only particles within these parameters were counted as strongyle eggs.

Strongyle eggs were counted in all samples using each method. All three methods were used to count each fecal slurry in quadruplicate. Therefore, each sample was counted a total of 12 times. All counting methods had the same multiplication factor of 20 for counting the entire chamber, and 66.67 for counting within the gridded portion only.

2.2. Precision of Mini-FLOTAC, McMaster, and smartphone system

This part of the study was to evaluate and compare the precision of the Mini-FLOTAC, a modified McMaster, and a revised prototype of the smartphone system. A new set of fecal samples were collected from horses naturally infected with mixed-species strongyle parasites. Strongyle egg counts were performed in triplicate, and each sample was counted using all three evaluated methods.

2.2.1. Horses

This portion of the study utilized 50 fecal samples collected in the same manner from the herd previously described, and some horses were collected from on more than one occasion. Samples were refrigerated and processed within 2 weeks of the collection date.

2.2.2. Egg counting methods

The three methods used to in this study section were the Mini-FLOTAC, the McMaster, and a more refined version of the smartphone prototype system. The samples were first prepared in the Fill-FLOTAC apparatus with glucose-NaCl as previously described. Each fecal slurry was used to perform triplicate counts with each method.

For the Mini-FLOTAC, the two 1-mL chambers on the Mini-FLOTAC slide were filled as described by Noel et al. (2017). The slides were allowed to rest for 10 min before counting the eggs under the microscope. For the McMaster, a single McMaster chamber was filled. The slide was allowed to rest for 10 min before counting under the microscope.

For the smartphone system, the samples were prepared slightly differently than previously described in order to improve sample processing and image analysis. First, a secondary filtering mechanism was constructed, consisting of a stack of three cell strainers with sequential pore sizes of 100 μm, 200 μm, and 400 μm placed atop a 50 mL conical tube, with the smallest filter attached to the manufacturer's collar. The fecal suspension was gently drawn through this filtering apparatus. Next, 3 mL of the egg-containing filtrate was pipetted onto another filter (20 μm), developed by Slusarewicz et al. (2016). This filter (with integrated Luer syringe port) was constructed to screw directly onto a 50 mL conical tube and was fitted with a small vibration motor which functioned to mix reagents and evenly spread the fecal material over the filter surface. One mL of a 2.4% bleach solution was pipetted onto the filter and incubated at room temperature for 30 s before being removed. Next, 1 mL of fluorescent staining agent was added and incubated at room temperature for 2 min, with the vibration turned on for the entire interval. Lastly, after removal of the stain, the sample was rinsed three times with PBS, and pulse vibrated for 1 s each time.

The 20 μm filter containing the prepared sample was unscrewed from the conical tube and screwed into the bottom of the smartphone docking unit, as described by Slusarewicz et al. (2016). The smartphone application was used to take an image of the filter and subsequently count the strongyle eggs using the settings described above.

Each method was used to perform triplicate slurry counts of strongyle eggs from all samples. Therefore, each sample was counted a total of 9 times. In this experiment, the Mini-FLOTAC had a multiplication factor of 5, the McMaster had a multiplication factor of 66.67, and the smartphone system had a multiplication factor of 3.83 EPG. The system Images 87% of the filter, and this is taken into account with the multiplication factor to correct for the proportion of the filter that is actually imaged.

2.3. Accuracy of Mini-FLOTAC, McMaster, and smartphone system

This study section utilized spiked fecal samples to observe and compare the accuracy of the Mini-FLOTAC (Barda et al., 2013), the McMaster (Roepstorff and Nansen, 1998), and the second prototype of the smartphone system as described above. The samples were spiked with levels of 5, 50, 500 or 1000 EPG, or left negative as a control. The samples were assigned with a random code, and performed in a randomized order. Analysts were blind to this information.

2.3.1. Horses

Fecal material was collected from a horse previously treated with an effective anthelmintic. Three repeated Mini-FLOTAC counts were used to confirm the fecal material was FEC negative. The strongyle eggs used for spiking were harvested from feces collected from an untreated and consistently high-egg shedding horse at the University of Kentucky research farm.

2.3.2. Isolation of strongyle eggs from feces

Strongyle eggs were harvested from fecal material using the same principle described by Noel et al. (2017).

2.3.3. Spiking of the samples

A total of 72 fecal samples were prepared per technique, where 36 of them remained negative to serve as negative controls, and the other 36 were spiked with a known number of strongyle eggs. The true spiked EPG levels tested were 5, 50, 500, and 1000 EPG. Each spike level was repeated 9 times. For each technique examined, a fecal sample of appropriate weight was obtained (i.e. 5 g for the Mini-FLOTAC and smartphone system and 4 g for the McMaster). Eggs were pipetted from the microtiter well-plate and onto the pre-weighed fecal sample. The well and pipette were checked under the microscope to ensure all eggs had been successfully transferred. Sample preparation continued where the Mini-FLOTAC and smartphone system samples were separately prepared as previously described in Section 2.2.2, and the McMaster samples were prepared as described by Roepstorff and Nansen (1998), resulting in a multiplication factor of 50. The 5 EPG spike level was omitted for the McMaster due to its detection limit of 50 EPG. This entire protocol was repeated three times.

2.4. Statistical analysis

2.4.1. Statistical analysis for precision

The coefficient of variation, defined by the standard deviation divided by the mean count times 100%, was calculated for each set of repeated counts for each sample per method. Mean coefficients of variation with 95% confidence intervals were then calculated for each of the three techniques. The percent precision for each technique was calculated by subtracting the coefficient of variation from 100 and was calculated for both counting the entire chamber as well as the volume underneath the grid of the McMaster slide. Linear correlations with the calculation of R^2 values were achieved using Microsoft Excel 2010 (Redmond, WA). Additional statistical analyses were performed using SAS software (version 9.4, SAS Institute, Cary, North Carolina, USA).

Here, the *t*-test procedure was used to evaluate presence of possible unequal variance between methods, with 'Method' as the categorical variable, and 'Total egg count' as a continuous variable. Lastly, a mixed linear model analysis was performed to observe possible differences in the magnitude of egg counts generated among the three different techniques. 'Horse ID,' 'Date,' and 'Replicate' were considered random variables. Whenever the variable 'Method' was found significant, a 'least squares means' analysis was used for a Tukey's pair-wise comparison ($\alpha = 0.05$).

2.4.2. Statistical analysis for accuracy, sensitivity, and specificity

The mean EPG for the three replicate counts were determined for each sample. Accuracy was calculated for each count using the following formula, where the expected counts are the true spiked EPG values, and the observed counts were the means of the triplicate EPGs.

$$\left(\frac{Observed}{Expected}\right) \times 100\%$$

The accuracy values for each spike level were then averaged to obtain a total percent accuracy for each method. Additional statistical analyses were performed using SAS software (version 9.4, SAS Institute, Cary, North Carolina, USA). A mixed linear model analysis was performed to evaluate and compare each techniques' accuracy. A deviation variable was created by subtracting the observed count from the true spiked count. Whenever this value was negative, the absolute value was used. 'Method' and 'True EPG' were assigned as categorical variables, and 'Replicate' as a random variable. When significant differences were found within 'Method' a least squares means analysis was used for a Tukey's pair-wise comparison ($\alpha = 0.05$). Lastly, diagnostic sensitivity, specificity, predictive values, and likelihood ratios were calculated using online software (MedCalc Software bvba, Ostend, Belgium). The accompanying 95% confidence intervals were examined to evaluate significant differences between techniques.

3. Results

3.1. Precision of computerized egg counting systems and McMaster

The coefficient of variation, mean of the standard deviations, precision estimates, and coefficient of determination (R^2) with corresponding *p*-values for counts of the entire McMaster chamber and within the grid generated by all three methods are presented in Table 1. The McMaster technique performed with significantly less variance than the two computerized methods ($p < 0.0001$). The smartphone and ImageJ performed with equal variance. A least squares means analysis found ImageJ to generate significantly higher egg counts than the smartphone system ($p < 0.0001$), and the smartphone system generated significantly higher egg counts than the McMaster ($p < 0.0001$).

Table 2
Precision evaluation of the Mini-FLOTAC, McMaster, and smartphone system (study section 2.2) represented by the coefficient of variation (CV), mean of standard deviation (SD), and precision estimates for strongyle egg counts. The detection sensitivity (EPG) for the McMaster was 66.67, the Mini-FLOTAC was 5, and the smartphone system was 3.83. 95% confidence intervals are included in parenthesis

Method	CV	Mean SD (EPG)	% Precision
McMaster	50.88 (36.68–65.08)	85.24	49.12
Mini-FLOTAC	35.66 (24.67–46.64)	43.33	64.34
Smartphone	28.4 (22.51–34.29)	58.52	71.6

Abbreviation: EPG, Eggs per gram.

3.2. Precision of Mini-FLOTAC, McMaster, and smartphone system

The coefficient of variation, mean of standard deviation, and percent precision of counts generated by the Mini-FLOTAC, McMaster, and smartphone system are shown in Table 2. The linear relationships between each of the three techniques and the corresponding coefficients of determination are presented in Fig. 1. Significantly unequal variances were found between all three methods, where the Mini-FLOTAC had significantly less variance than the McMaster ($p = 0.0228$), and the smartphone system had significantly less variance than the Mini-FLOTAC ($p < 0.0001$) and McMaster ($p < 0.0001$). The mixed model analysis did not identify any significant differences in the magnitude of strongyle egg counts between methods ($p = 0.4802$).

3.3. Accuracy of Mini-FLOTAC, McMaster, and the smartphone system

A graphical representation for the observed egg counts compared to the true egg counts for each method is shown in Fig. 2. Accuracy estimates for the Mini-FLOTAC, McMaster, and the smartphone system at each spike level (EPG) are shown in Table 3. The Mini-FLOTAC was significantly more accurate than the McMaster ($p < 0.0001$) and the smartphone system ($p < 0.0001$). There was no significant difference between the accuracy of the smartphone system and the McMaster technique ($p = 0.9411$). A significant difference was found between the 500 and 1000 EPG spike levels for the Mini-FLOTAC ($p < 0.0001$) and the smartphone system ($p = 0.0024$).

Table 1
Precision evaluation of the McMaster, Image J, and smartphone system strongyle egg counting methods (study section 2.1) represented by the coefficient of variation (CV), mean of standard deviation, and precision estimates. 95% confidence intervals are included in parenthesis. Alsoshown are the coefficients of determination (R^2) between methods with the corresponding p-values for strongyle egg counts. All counting methods had the same multiplication factor of 20 for counting the entire chamber, and 66.67 for counting within the gridded portion only.

	Entire Chamber[a]			Grid Only[b]		
Method	CV	mean of SD (EPG)	% precision	CV	mean of SD (EPG)	% precision
McMaster	21.14% (6.96–35.32)	165.98	78.86	18.94 (14.76–23.12)	248.68	81.06
ImageJ	30.61% (25.05–36.17)	1025.83	69.39	31.79 (26.07–37.51)	1253.40	68.21
Smartphone	32.79% (26.86–38.72)	836.57	67.21	37.08 (30.57–43.60)	1560.54	62.92

Coefficients of determination

Method	R^2	p-value	R^2	p-value
Smartphone vs McMaster	0.14	0.080	0.19	0.044
Smartphone vs ImageJ	0.29	0.014	0.44	< 0.001
McMaster vs ImageJ	0.30	0.008	0.31	0.008

Abbreviation: EPG, Eggs per gram.
[a] Entire chamber represents eggs counted in the entire McMaster chamber.
[b] Grid only represents eggs count in the McMaster grid only.

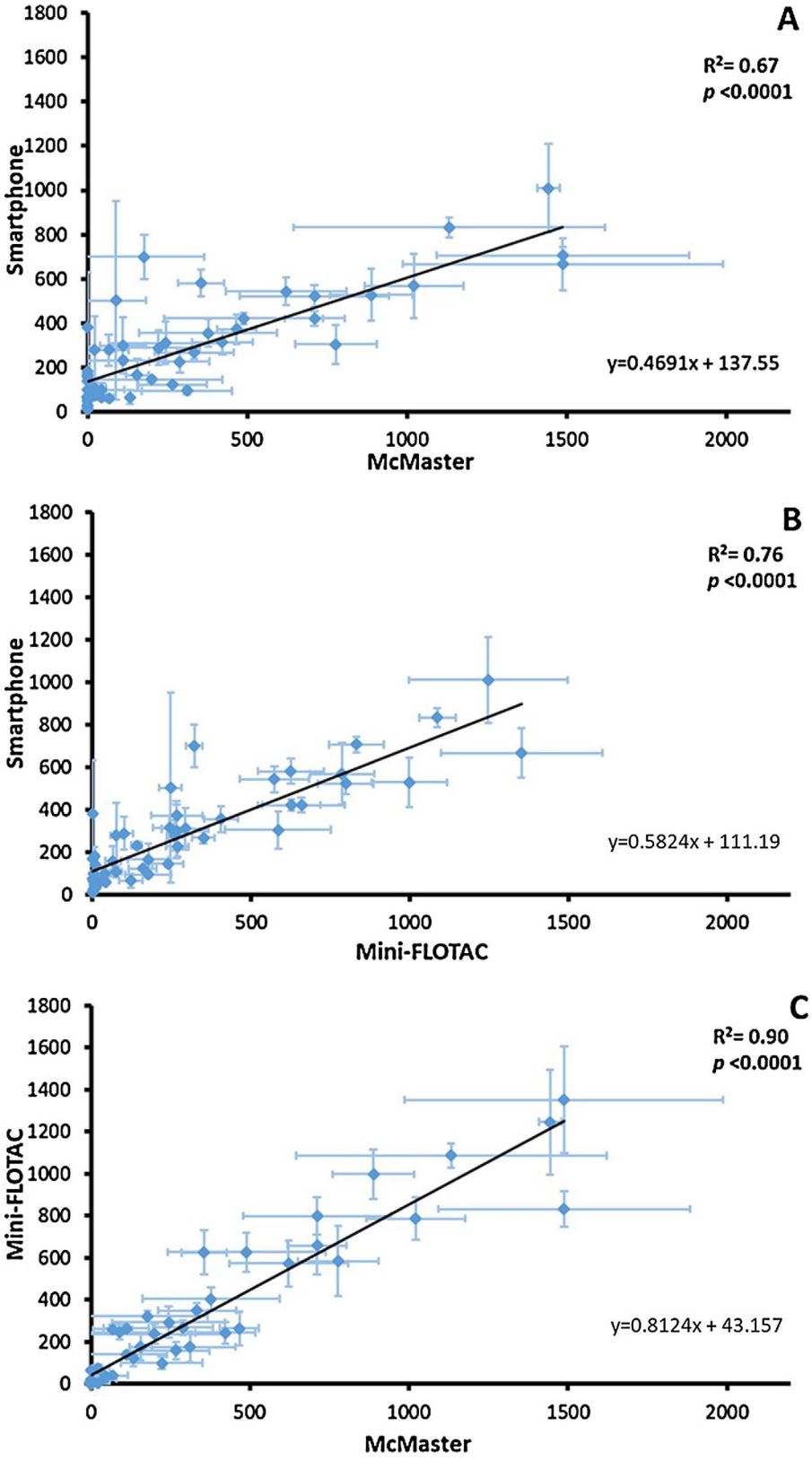

Fig. 1. Scatter plots comparing strongyle fecal egg counts in eggs per gram for the smartphone system and McMaster (A), smartphone system and Mini-FLOTAC (B), and Mini-FLOTAC and McMaster (C). Error bars designate the 95% confidence interval between repeated counts on each sample.

3.4. Sensitivity, specificity, predictive values, and likelihood ratios for the Mini-FLOTAC, McMaster and the smartphone system

Sensitivity, specificity, predictive values, and likelihood ratios for each method are shown in Table 4.

4. Discussion

The aim of this study was to compare an image analysis based egg counting method with manual methods. The refinements of the second smartphone system prototype reduced operator dependency. As a result, the smartphone system had significantly better precision than both

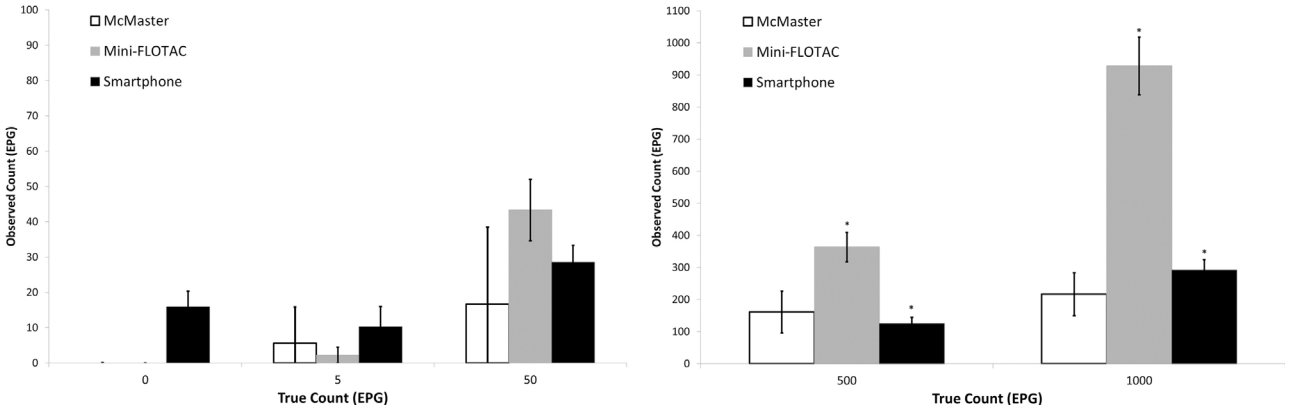

Fig. 2. Graphical representation of the observed strongyle egg counts in eggs per gram (EPG) generated by the McMaster (white), Mini-FLOTAC (grey), and smartphone system (black) in comparison to the true spiked count. Error bars represent 95% confidence intervals, and the asterisks indicate statistically significant differences between methods within spike levels ($\alpha = 0.05$).

the McMaster and Mini-FLOTAC techniques. It was, however, significantly less accurate than the Mini-FLOTAC and had comparable accuracy to the McMaster. The results indicate that the smartphone system prototype has limitations and further refinement should be aimed at improving accuracy, sensitivity, and specificity.

Inconsistent egg identification of the egg counting algorithms, such as inconsistent assignment of non-egg particles as eggs, resulted in lower precision of the smartphone system than the McMaster in the first study section. At this stage, the algorithm also could not readily differentiate two eggs in close proximity or egg stacking. Furthermore, in this initial setup, eggs were manually transferred from the filter to the McMaster counting chamber, which is likely to have contributed considerably to the variability.

This study also provided a comparison of precision between eggs counted in the entire McMaster chamber and eggs counted solely within the gridded portion (the more commonly used method in clinical practice). It is often presumed that a larger examined sample volume will result in better precision (Lester and Matthews, 2014). Several studies showed that examining larger sample volumes than the typical McMaster can result in greater precision (Rinaldi et al., 2011; Levecke et al., 2012; Bosco et al., 2014). Interestingly, our findings did not support these conclusions as we did not find significant differences in

Table 3
Percent accuracy of the McMaster, Mini-FLOTAC and smartphone system (study section 2.3) is shown for the three methods at each spike level (EPG). The total accuracy represents the average accuracy for each method across all spike levels. 95% confidence intervals are included in parenthesis. Accuracy = (observe/expected)x100%

Method (sensitivity level)	5 EPG (%)	50 EPG (%)	500 EPG (%)	1000 EPG (%)	Total (%)
McMaster (50 EPG)	NA	11.11	32.22	21.67	21.67 (13.87–29.47)
Mini-FLOTAC (5 EPG)	22.22	75.56	72.33	87.94	64.51 (53.19–75.84)
Smartphone (3.83 EPG)	18.91	57.02	25.02	29.15	32.53 (25.29–39.76)

Abbreviation: EPG, Eggs per gram; NA, not applicable.

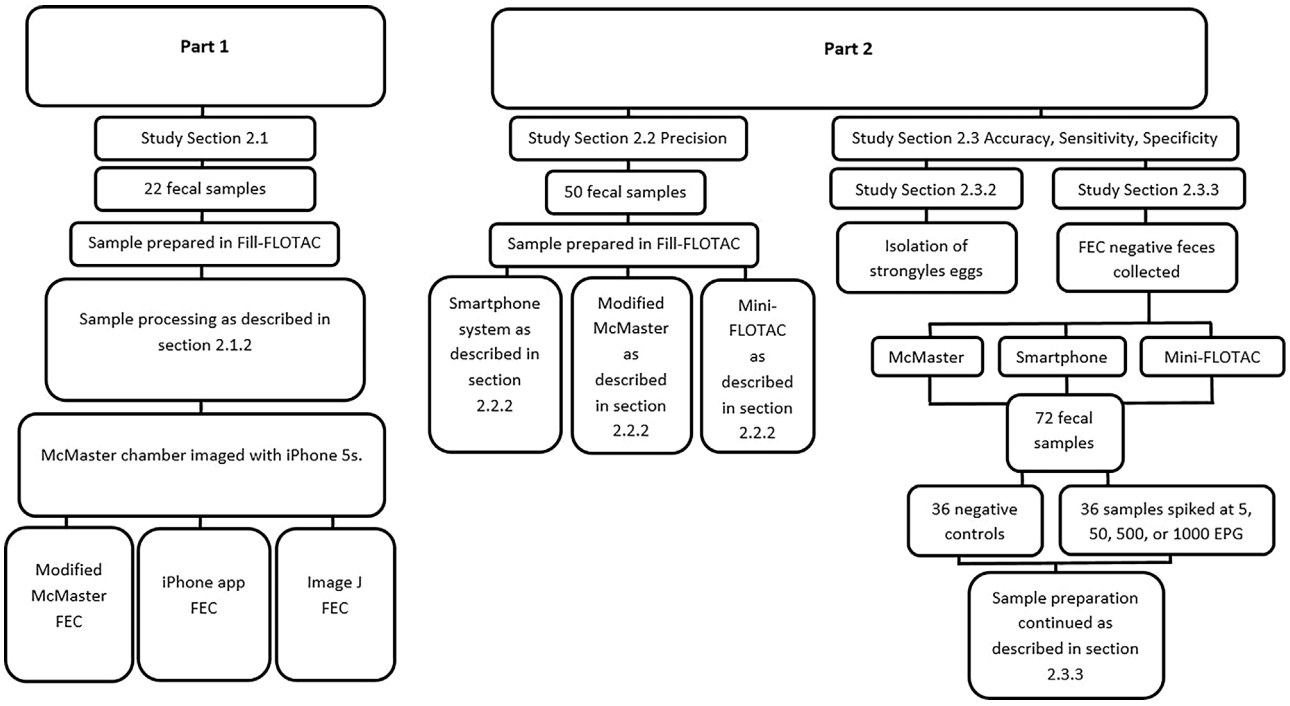

Fig. 3. Flow chart representation of the study design and methodology.

Table 4
Sensitivity, specificity, predictive values, and likelihood ratios for fecal egg counts performed with the Mini-FLOTAC, the smartphone system, and McMaster methods (study section 2.3). 95% confidence intervals are included in parenthesis.

Diagnostic Parameter	Mini-FLOTAC	Smartphone	McMaster
Sensitivity	83.33% (67.19–93.63)	94% (81.34–99.32)	55.56% (38.1–72.06)
Specificity	100% (90.26–100)	14% (4.67–29.5)	100% (90.26–100)
Positive Predictive Value	100% (88.43–100)	52% (39.54–64.85)	100% (83.16–100)
Negative Predictive Value	85.71% (71.46–94.57)	71% (29.04–96.33)	69.23% (54.9–81.28)
Positive Likelihood Ratio	NA	1.10 (0.94–1.28)	31.67 (4.49–223.35)
Negative Likelihood Ratio	0.17 (0.08–0.35)	0.40 (0.08–1.93)	0.3 (0.17–0.54)

Abbreviation: NA, not applicable.

precision between the two sample volumes examined, 0.15 and 0.5 mL. One possible explanation for this is an uneven distribution of eggs within the McMaster chamber, as proposed by Cringoli et al. (2004). While the aforementioned study did examine various sample volumes, the study design also included very different sample preparation steps including creating different sample dilutions, and it is not possible to make direct comparisons between studies. It is possible that the differences in volumes we examined may not have been large enough to produce a difference in precision, but it is also possible that some of the differences described by others may have been due, at least partially, to differences in other parts of the methodologies (most notably sample homogenization). Therefore, while increasing the sample volume examined can improve precision, it may not be the only affecting factor.

The second prototype of the smartphone system resulted in several refinements, particularly to reduce operator dependency. This prototype had significantly better precision among repeated counts than the McMaster and Mini-FLOTAC. It is important to note that only one chamber of the McMaster slide was examined, resulting in a high limit of detection (66.67) and is a limitation of this study. However, as the calculation of CV is not affected by multiplication factor (limit of detection), this study did provide evidence that the smartphone system performed with better precision than the two manual methods. The opportunity for egg loss and sample variability was likely reduced by performing all sample preparation steps on a single filter. The filter could also be directly attached to the unique docking unit for imaging, eliminating the need for harvesting the sample from the filter for placement into the McMaster chamber for imaging. This reduction in operator dependency and improvements in sample preparation may explain the increased precision for the smartphone system. However, the filtration step likely contributed a source of error as it can result in egg stacking and/or eggs can be fully or partially covered by debris, causing the algorithm to not include them in the egg count. The system presented here is a prototype and refinement is ongoing.

Concerning accuracy, the Mini-FLOTAC method was significantly more accurate than the McMaster and smartphone system. This suggests that the Fill-FLOTAC is an efficient way for homogenizing the sample and distributing it directly into the Mini-FLOTAC slides, reducing the opportunity for egg loss in comparison to the sample preparation described for the McMaster and smartphone system. The better accuracy of the Mini-FLOTAC over the McMaster is in agreement with recent results generated in our laboratory (Noel et al., 2017). The different detection limits between methods will largely affect these values, and the McMaster method performed in this study section had a detection limit of 50 EPG and cannot be expected to detect eggs below this level. Furthermore, we found statistically significant differences in egg counts between the 500 and 1000 EPG spike levels for the Mini-FLOTAC and smartphone system. Given the spiked egg count was doubled, the observed egg counts should also double in response. In contrast, the McMaster method did not exhibit a statistically significant difference between these levels. In regards to sensitivity, the smartphone system was significantly more sensitive than the McMaster, while the Mini-FLOTAC was not significantly different from the other two methods. The smartphone system had significantly lower specificity than the other two methods, and the Mini-FLOTAC and McMaster were not significantly different from each other. These results are likely due to the tendency of the smartphone system to generate false-positive egg counts, resulting in a higher sensitivity and lower specificity. This warrants further improvement of the algorithm for detecting low egg count samples which are particularly important for performing the fecal egg count reduction test.

Interestingly, counts performed on the naturally infected samples revealed better agreement between the McMaster and Mini-FLOTAC methods than that between the McMaster and smartphone system. In contrast, the McMaster and smartphone system exhibited more similar egg recovery abilities from spiked samples than the McMaster and Mini-FLOTAC. These contradictory results suggest that the spiked samples only represent a simulation of naturally infected samples and therefore may not truly represent the techniques' egg recovery capabilities. Another observation was the variation between slopes observed in the scatter plots between the different techniques (Fig. 1), which illustrates that the range of FECs detected by these techniques will likely be different due to systematic differences between methods and that a procedure for standardization of FECs across techniques should be encouraged.

Overall, this smartphone-based automated egg counting system was more precise than the Mini-FLOTAC and McMaster methods, as well as comparable in accuracy to the McMaster technique, the most commonly used technique in clinical practice. Further refinement of the algorithm will likely further improve accuracy, while reducing operator dependency during sample preparation will provide a more user-friendly and standardized egg counting method. This method may offer improvement for surveillance-based equine deworming programs in the form of a precise on-site, standardized, automated FEC method, however further development of the technology is still needed.

Conflict of interest statement

PS and MKN both hold stock in MEP Equine Solutions, LLC and PS is an employee of this company, which is seeking to commercialize a mobile smartphone-based parasite egg counting technology for equine usage. No other authors have any conflicts of interest to declare.

Acknowledgements

This work was funded by a Small Business Innovation Research Grant from the United States Department of Agriculture (#2015-33610-23497). The authors are very grateful to MEP Equine Solutions, LLC and Eric Hauck for providing the materials, resources and appropriate training for the smartphone system. We also extend our gratitude to Chris Mills and Stephanie Pagano for assisting with the development of the smartphone system.

References

Ballweber, L.R., Beugnet, F., Marchiondo, A.A., Payne, P.A., 2014. American Association of Veterinary parasitologists' review of veterinary fecal flotation methods and factors

influencing their accuracy and use-Is there really one best technique? Vet. Parasitol. 204, 73–80.
Barda, B.D., Rinaldi, L., Ianniello, D., Zepherine, H., Salvo, F., Sadutshang, T., Cringoli, G., Clementi, M., Albonico, M., 2013. Mini-FLOTAC an innovative direct diagnostic technique for intestinal parasitic infections: experience from the field. PLOS Negl. Trop. Dis. 7, 1–7.
Bosco, A., Rinaldi, L., Maurelli, M.P., Musella, V., Coles, G.C., Cringoli, G., 2014. The comparison of FLOTAC, FECPAK and McMaster techniques for nematode egg counts in cattle. Acta Parasitol. 59, 625–628.
Carstensen, H., Larsen, L., Ritz, C., Nielsen, M.K., 2013. Daily variability of strongyle fecal egg counts in horses. J. Equine Vet. Sci. 33, 161–164.
Castañón, C.A.B., Fraga, J.S., Fernandez, S., Gruber, A., Costa, L.F., 2007. Biological shape characterization for automatic image recognition and diagnosis of protozoan parasites of the genus *Eimeria*. Pattern Recogn. 40, 1899–1910.
Chapman, M.R., French, D.D., Klei, T.R., 2003. Prevalence of strongyle nematodes in naturally infected ponies of different ages and during different seasons of the year in Louisiana. J. Parasitol. 89, 309–314.
Cringoli, G., Rinaldi, L., Veneziano, V., Capelli, G., Scala, A., 2004. The influence of flotation solution, sample dilution and the choice of McMaster slide area (volume) on the reliability of the McMaster technique in estimating the faecal egg counts of gastrointestinal strongyles and *Dicrocoelium dendriticum* in sheep. Vet. Parasitol. 123, 121–131.
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.
Dogantekin, E., Yilmaz, M., Dogantekin, A., Avci, E., Sengur, A., 2008. A robust technique based on invariant moments − ANFIS for recognition of human parasite eggs in microscopic images. Expert Syst. Appl. 35, 728–738.
Egwang, T.G., Slocombe, J.O.D., 1982. Evaluation of the Cornell-Wisconsin centrifugal flotation technique for recovering trichostrongylid eggs from bovine feces. Can. J. Comp. Med. 46, 133–137.
Ghazali, K.H., Hadi, R.S., Mohamed, Z., 2013. Automated system for diagnosis intestinal parasites by computerized image analysis. Mod. Appl. Sci. 7, 98–114.
Godber, O.F., Phythian, C.J., Bosco, A., Ianniello, D., Coles, G., Rinaldi, L., Cringoli, G., 2015. A comparison of the FECPAK and Mini-FLOTAC faecal egg counting techniques. Vet. Parasitol. 207, 342–345.
Grishagin, I.V., 2015. Automatic cell counting with Image. J. Anal. Biochem. 15, 63–65.
Kuzmina, T.A., 2012. Strongylids (Nematoda: Strongylidae) of domestic horses in Ukraine: modern state of fauna and structure of the parasite community. Parazitolgiya 46, 127–138.
Lester, H.E., Matthews, J.B., 2014. Faecal worm egg count analysis for targeting anthelmintic treatment in horses: points to consider. Equine Vet. J. 46, 139–145.
Levecke, B., Rinaldi, L., Charlier, J., Maurelli, M.P., Bosco, A., Vercruysse, J., Cringoli, G., 2012. 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.
MAFF (Ministry of Agriculture, Fisheries and Food), 1986. Manual of Veterinary Parasitological Laboratory Techniques, 3rd ed. HMSO, London 24 pp.
McCoy, M.A., Edgar, H.W.J., Kenny, J., Gordon, A.W., Dawson, L.E.R., Carson, A.F., 2005. Evaluation of on-farm faecal worm egg counting in sheep. Vet. Rec. 156, 21–23.
Mes, T.H.M., Eysker, M., Ploeger, H.W., 2007. A simple, robust and semi-automated parasite egg isolation protocol. Nat. Prot. 2, 486–489.
Nielsen, M.K., Mittel, L., Grice, A., Erskine, M., Graves, E., Vaala, W., Tully, R.C., French, D.D., Bowman, R., Kaplan, R.M., 2016. AAEP Parasite Control Guidelines. American Association of Equine Practitioners Online at www.aaep.org on 2/1/2017.
Noel, M.L., Scare, J.A., Bellaw, J.L., Nielsen, M.K., 2017. Accuracy and Precision of mini-FLOTAC and McMaster techniques for determining equine strongyle egg counts. J. Equine Vet. Sci. 48, 1–6.
Rinaldi, L., Coles, G.C., Maurelli, M.P., Musella, V., Cringoli, G.C., 2011. Calibration and diagnostic accuracy of simple flotation: McMaster and FLOTAC for parasite egg counts in sheep. Vet. Parasitol. 177, 345–352.
Roepstorff, A., Nansen, P., 1998. Epidemiology, diagnosis, and control of helminth parasites of swine. FAO Animal Health Manual No 3. Food and Agriculture Organization of the United Nations, Rome.
Slusarewicz, P., Pagano, S., Mills, C., Popa, G., Chow, K.M., Mendenhall, M., Rodgers, D.W., Nielsen, M.K., 2016. Automated parasite faecal egg counting using fluorescence labelling, smartphone image capture and computational image analysis. Int. J. Parasitol. 46, 285–493.
Stoll, N., 1930. On methods of counting nematode ova in sheep dung. Parasitology 22, 116–136.
Suzuki, C.T.N., Gomes, J.F., Falcão, A.X., Papa, J.P., Hoshino-Shimizu, S., 2013. Automatic segmentation and classification of human intestinal parasites from microscopy images. IEEE T. Bio-med. Eng. 60, 803–812.
Vidyashankar, A.N., Hanlon, B.M., Kaplan, R.M., 2012. Statistical and biological considerations in evaluating drug efficacy in equine strongyle parasites using fecal egg count data. Vet. Parasitol. 185, 45–46.
Yang, Y.S., Park, D.K., Kim, H.C., 2001. Automatic identification of human helminth eggs on microscopic fecal specimens using digital image processing and an artificial neural network. IEEE Trans. Bio-med. Eng. 48, 718–730.