Short communication

Intra- and inter-observer agreement when using a descriptive classification scale for clinical assessment of faecal consistency in growing pigs

Ken Steen Pedersen *, Nils Toft

Department of Large Animal Sciences, University of Copenhagen, Groennegaardsvej 2, DK-1870 Frederiksberg C, Denmark

ARTICLE INFO

ABSTRACT

The objective of the current study was to evaluate intra- and inter-observer agreement using a descriptive classification scale with four categories, descriptive text and pictures for assessment of consistency in faecal samples from pigs post weaning. The four consistency categories were score one = firm and shaped, score two = soft and shaped, score three = loose and score four = watery.

Five observers from the same veterinary practice examined 100 faecal samples using the scale with four categories. Four of the observers examined the 100 faecal samples twice within the same day.

Within observers the difference in proportions for the individual consistency categories between two examinations was on average 0.04 (range: 0–0.10). The mean intra-observer agreement was 0.82 (range: 0.72–0.91) with a mean kappa value of 0.76 (range: 0.61–0.88).

For inter-observer agreement overall kappa was 0.64. For the 10 pair-wise comparisons the mean inter-observer agreement was 0.73 (range: 0.61–0.90) with a mean kappa value of 0.64 (range: 0.48–0.87). The difference in proportions for the individual consistency categories was on average 0.08 (range: 0–0.17).

In conclusion, the agreement observed for the descriptive classification scale with four categories, descriptive text and pictures may be categorized as a substantial to almost perfect intra-observer agreement and a moderate to almost perfect inter-observer agreement.

However, more objective measures than clinical scales may still be needed to improve intra- and inter-observer agreement in research studies.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Diarrhoea in pigs post-weaning continues to be important in countries around the world. Diarrhoea has been defined as an increased frequency of defecation accompanied by faeces which contain an increased concentration of water and decreased dry matter content (Radostits et al., 2000). Clinical evaluation of the consistency of faeces and hence classification of pigs with diarrhoea is performed in both research and veterinary practice. There is no standardized protocol for clinical characterization of faecal consistency. This makes comparison among various diarrhoea studies difficult. Further, clinical assessment of faecal consistency is partly subjective. Variation within and between observers in clinical assessment of faecal consistency affects the ability to obtain valid and repeatable data in research and veterinary practice. A large variation in inter-observer agreement has been reported when trying to classify faecal samples from pigs as normal, loose and fluid (Pedersen et al., 2011). However, a good inter-observer agreement has been reported in human medicine using a descriptive classification scale

* Corresponding author. Tel.: +45 3533 3068; fax: +45 3533 3022.
E-mail address: ken@life.ku.dk (K.S. Pedersen).

for assessment of faecal consistency (Bliss et al., 1999, 2001). Similar results may be obtained using a descriptive classification scale for assessment of faecal consistency in pigs.

The objective of the current study was to evaluate intra- and inter-observer agreement using a descriptive classification scale with four categories, descriptive text and pictures for assessment of consistency in faecal samples from pigs post weaning.

2. Materials and methods

2.1. Faecal classification scale

A classification scale with four descriptive categories and explanations in text and pictures for assessment of faecal consistency was developed by observer one. The classification scale was modified compared to the previous evaluated classification scale with three categories (Pedersen et al., 2011). The category normal was divided in two categories (score one and two) in order to provide a better description of faecal samples with normal consistencies.

The classification scale with associated text and pictures is illustrated in Fig. 1. The classification scale was developed for clinical assessment of faecal samples on the pen floor as well as examination of faecal samples in plastic containers.

2.2. Sample size considerations

The results reported by Pedersen et al. (2011) were used for sample size calculations. Preliminary sample size considerations were based on expectations for intra-observer and pair wise inter-observer agreement using formulae for estimating proportions. The CII command in Stata IC version 11 was used. Using an expected agreement of 0.70, a sample size of 100 faecal samples would provide an acceptable 95% confidence interval (allowable error of approximately 0.10) for the intra- and inter-observer agreement estimates.

Samples size considerations were further based on expected kappa values in relation to pair wise inter-observer agreement. The SSKDLG command in Stata IC version 11, based on procedure outlined by Cantor (1996), was used to determine sample size for a range of different kappa values (0–0.95) and proportions of positive samples by each observer (0.1–0.9). A 95% confidence interval with an allowable error of 0.20 for kappa was used in the calculations and a sample size of 100 faecal samples was determined to be sufficient.

2.3. Observers

Five observers were selected for the study of inter-observer agreement due to the practical limitation of completing all examinations within one day. Only four observers were able to participate in intra-observer agreement testing for the same reason. It was intended to mimic a best case scenario without inter-observer calibration. A convenience sample of five observers from the same specialized swine practice was selected to avoid differences due to practice specific conditions. These five observers included the corresponding author, observer one.

2.4. Evaluation of intra- and inter-observer agreement

A diagram (Fig. 1) with explanations in text and pictures of faeces representing each of the four consistency categories was send by e-mail to the observers four days prior to examination. The observers were instructed to familiarise themselves with the classification scale prior to the examination.
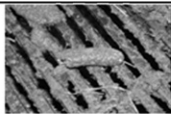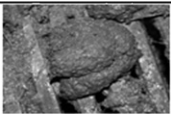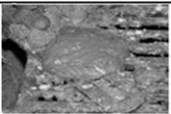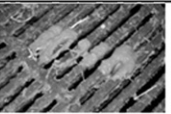
Score	1 Firm and shaped	2 Soft and shaped	3 Loose	4 Watery
Picture				
Texture	Firm. Varies in hardness.	Varies in softness. Like peanut butter	Mush. Often shining surface	Varies form gruel to water.
Shape	Sausage	Varies form sausage shape to small piles	Tends to level with surface. Does not flow through or flows slowly through slatted floors.	Levels with surface. Flows through slatted floors.
In container	Preserves original shape.	Does not flow when container is rotated. Preserves original shape.	Inert when container is rotated. Merges and covers bottom of container in most cases.	Flows easy when container is rotated. Merges and covers bottom of container.

Fig. 1. Faecal classification scale with 4 categories, descriptive text and pictures.