Summarizing the evidence of the performance of diagnostic tests for *Salmonella* spp. in swine: A systematic review approach.

1 Wilkins, W.*, 2,5 Rajić, A., 1 Parker, S., 2,5 Waddell, L., 2,4 Sanchez, J., 2,3 Sargeant, J., 1 Waldner, C.

1 Western College of Veterinary Medicine, Saskatoon, Saskatchewan, Canada; 2 Public Health Agency of Canada, Guelph, Ontario, Canada; 3 McMaster University, Hamilton, Ontario, Canada; 4 Atlantic Veterinary College, Charlottetown, Prince Edward Island, Canada; 5 Ontario Veterinary College, Guelph, Ontario, Canada

*corresponding author: Wendy.Wilkins@usask.ca

Abstract

A systematic review was undertaken in order to identify, critically appraise and synthesize the existing literature on the diagnostic performance (e.g. Se and Sp of the test) and agreement (e.g. kappa and correlation coefficients) of conventional bacterial culture, enzyme-linked immunosorbent assays and polymerase chain reaction assays used to detect and monitor *Salmonella* spp. in swine. 2110 citations were identified and 160 were relevant to the research objectives. Quality assessment is complete for 150 of these references; 73 were excluded due to estimates of test performance not being clearly reported and insufficient raw data available for post-hoc analysis. Although the review is still in progress, early experiences indicate problems with lack of standardization in the design, conduct and reporting of studies of diagnostic test evaluation in this area. This review will provide valuable information by identifying gaps in existing research and providing direction for future work on the standardization of tests examined in this review.

Introduction

Recent initiation or consideration of *Salmonella* control programs in swine in many European pig-producing countries has created an impetus for other pork producing countries to investigate the epidemiology of *Salmonella* infection in their pig populations and to evaluate the feasibility of potential control options. Although the Danish surveillance program for *Salmonella* in swine has been in place for almost a decade (Mousing et al. 1997), it has been only recently that the international research community has started to address some basic research questions related to the validity and accuracy of existing sampling strategies and testing protocols, particularly at the farm level.

A systematic review is an overview of primary studies which contains an explicit statement of objectives, materials, and methods and has been conducted according to explicit and reproducible methodology (Greenhalgh 1997). This methodology has been increasingly used in human medicine to synthesize the results of studies of diagnostic test accuracy; however, quality systematic reviews in general in the areas of animal health and agri-food public health are limited (Sargeant et al. 2006). The application of systematic reviews in these areas is complicated by the use of challenge studies and observational studies, as well as the lack of randomized controlled trials (Sargeant et al. 2006). While these factors present an additional challenge to researchers, the obstacle is not insurmountable. Through the process of addressing the review question the researchers may also set precedence for review protocol.

In human health and medicine, diagnoses are made to predict prognosis and to guide treatment decisions. In animal health and agri-food public health, diagnostic tests are more often used to evaluate herd prevalence and to classify herds for monitoring and control programs. The main objectives of this systematic review were to identify, critically appraise and synthesize the existing literature on the diagnostic performance (e.g. Se and Sp of the test) and agreement (e.g. kappa and correlation coefficients) of conventional bacterial culture, enzyme-linked immunosorbent assays and polymerase chain reaction assays used to detect and monitor *Salmonella* spp. in swine.
Material and methods

The initial objective of the literature search was to identify all publications reporting the evaluation of one or more tests used to detect *Salmonella* infection in domestic swine. In addition, we sought to identify all studies where two or more of these tests had been used simultaneously on the same subject/sample to detect *Salmonella*, recognizing that these studies potentially contain data that could permit the evaluation/comparison of the tests being used even though test evaluation was not one of the objectives of the research being reported. Thus a broad search was performed, designed to have a high sensitivity for any abstract reporting the use of a diagnostic test for *Salmonella* in swine. The databases searched, search terms and results for each database were recorded. The search was restricted to 1980 and onward, as it was thought that tests after this date would be most representative of tests in current use.

Two reviewers independently reviewed all references at each step of the review process. First, the titles and abstracts of all articles were screened to identify potentially relevant articles. Articles reporting the evaluation of one or more tests (ELISA/serology, bacterial culture or PCR) were retained for further evaluation, as were articles reporting the use of two or more of these tests to detect the presence of *Salmonella* in swine. These abstracts were then subject to a 2nd level relevance screening, where non-English references and non-primary research articles were excluded.

Next, quality assessment was done using the full texts of articles passing through relevance screening. Questions pertaining to study quality were restricted to those items which the review team deemed to be critical inclusion criteria, in order to expedite the review process. Other non-critical questions pertaining to study quality were included in the subsequent data extraction process. Finally, data was extracted from references surviving the quality assessment phase. The data extracted included general study information (including non-critical study quality questions), details of test protocols and reported measures of test performance or agreement, and/or raw test data, if available.

The review process was carried out using an electronic systematic review (eSR) program developed by TrialStat® (www.trialstat.com). All review forms were developed a priori and pre-tested prior to use. Disagreements were resolved by discussion between reviewers until consensus was reached.

Results

A total of 2110 citations were identified, uploaded to the eSR database and screened for relevance. 160 of these references were deemed relevant to the research objectives. Quality assessment is nearly complete, with screening of 10 references still in progress.

Of the 150 references for which quality assessment is complete, 73 have been excluded due to estimates of test performance not being clearly reported and insufficient raw data available for post-hoc analysis. Preliminary exploration of the data indicates that approximately 50% of research reporting the evaluation of one or more tests used to detect *Salmonella* infection in domestic swine did not report actual estimates of test performance or agreement, nor did they provide raw data in a manner which permits post-hoc analysis. Similarly, approximately 50% of the studies which were included because two or more of these tests had been used simultaneously on the same subject/sample contained no useable data with which to perform post-hoc analysis.

The remaining 77 references are in the process of data extraction. While there are insufficient data to explore at this time, early experiences point to potential difficulties with future data synthesis attempts. Some of the potential problems include: insufficient/no detail on test protocol(s); inconsistent use of reference tests; pooling of results from different populations or from different tests; and failure to specify the population from which samples were obtained.

Discussion

Systematic reviews evaluating diagnostic tests important in veterinary and agri-food public health are virtually non-existent in the published literature, despite the fact that review methodology in the human health fields is well developed. To illustrate, a simple search of the PubMed database, using the search string "systematic review AND diagnostic test", will return 52 results (as of
February, 2007); in contrast, using the same search string and combining it with veterinary or animal related terms will return no results. This failure to utilize systematic reviews in non-traditional areas is regrettable, as the use of systematic reviews to synthesize the current body of knowledge on targeted food safety issues – in this case diagnostic test performance - can provide increased credibility to findings in the field (Sargeant et al. 2005). The principles that apply to evaluating diagnostic tests in human health also apply in animal and agri-food public health; therefore, it is possible to adapt existing tools for systematic reviews of diagnostic tests and modify these tools for application in these areas. One such tool is the QUADAS tool (Whiting et al. 2003), which provides criteria for assessing the quality of studies of diagnostic test evaluation. As mentioned previously, the use of challenge studies and observational studies, and the lack of randomized controlled trials present a challenge to reviewers in veterinary and food safety fields(Sargeant et al. 2006), and subsequent modifications to tools such as the QUADAS tool must take this into consideration.

A major variation to the diagnostic test review protocol that was made in this review was the decision to include all studies where two or more of these tests had been used simultaneously on the same subject/sample to detect Salmonella, even though test evaluation was not an objective of these studies. For those studies that report the results of each test in a manner that allows extraction of this information into a two-by-two contingency table, estimates of test performance (percent agreement, kappa, sensitivity, specificity) can be calculated. This methodology may be particularly useful in cases where there is a scarcity of published studies regarding the performance of a diagnostic test.

Another variation to the more traditional systematic review protocol that we made was to include studies of all levels of evidence. In the human health field, studies included in systematic reviews are typically of the highest level of evidence – randomized control trials which are published in peer-reviewed journals (ref). In our review, studies of any design were included if they contained information relevant to the review question. Studies from ‘grey-literature’ sources (non-published research, e.g. conference proceedings) were also included, in contrast to more traditional systematic reviews. The impact of study design and literature type will be examined once data extraction is complete.

Early experiences with data extraction have hinted at potential problems with future data synthesis attempts. Insufficient detail of test protocol, a wide range of potential reference tests, pooling of results from different populations or from different tests, and failure to specify the population from which samples were obtained are examples of some of the problems encountered so far. Many of these types of problems are due to a lack of standardization in the design, conduct and reporting of studies of diagnostic test evaluation. This lack of standardization has been a problematic in the human health fields as well, and efforts have been made to encourage the research community to use a more structured approach (Meyer 2003). The anticipated outcome of the current systematic review is to perform a meta-analysis to calculate summary estimates of the diagnostic performance and agreement of these diagnostic tests; however if insufficient data are available, qualitative systematic review will still provide valuable information by identifying gaps in existing research and providing direction for future work on the standardization of tests examined in this review.

Conclusions

Systematic reviews are under-utilized in animal health and agri-food public health, and systematic reviews of diagnostic tests are virtually non-existent in these areas. The use of these tools in non-traditional areas is encouraged, as the use of systematic reviews to synthesize the current body of knowledge can provide increased credibility to findings in these fields, in addition to identifying gaps in existing research and providing direction for future work.

Acknowledgments

We would like to thank Janet Harris and Lindsay Downey for their contributions to this project. This project was funded by the Public Health Agency of Canada.
References

Greenhalgh T. Papers that summarise other papers (systematic reviews and meta-analyses). BMJ 1997; 315 (7109):672-5.


