A Guide to Conducting Systematic Reviews in Agri-Food Public Health

Jan M. Sargeant [1,2], Maria Del Rocio Amezcua [1,2], Andrijana Rajić[2], Lisa Waddell [2]

[1] Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada
and
[2]Laboratory for Foodborne Zoonoses, Public Health Agency of Canada, Guelph, Ontario, Canada

© J.M. Sargeant, 2005

Table of Contents

Funding for this manual was obtained from the Food Safety Research and Response Network (USDA-CREES) and the Laboratory for Foodborne Zoonoses, Public Health Agency of Canada.

Introduction

Food safety policy makers and other decision-makers in the food production continuum are continuously and increasingly dealing with complex multi-factorial public health issues such as foodborne illness, existing and emerging zoonotic pathogens, and antimicrobial resistance. For instance, foodborne illnesses cause an estimated 76 million illnesses and 5000 deaths annually in the United States (Mead, 1999). Many organizations and agencies have addressed these concerns through an increased scope and intensity of targeted surveillance, the development and implementation of intervention strategies throughout the food chain, the integration of risk assessment into policy development and through increased funding to food safety research.

There is a need to start identifying, appraising and summarizing the results of otherwise unmanageable quantities of agri-food public health research so that decision-makers can access timely information on the most relevant scientific literature. This can be accomplished by utilizing evidence-based systematic review methodologies that have been used successfully in other health disciplines.

Systematic reviews provide a rigorous and replicable method of identifying, evaluating, and summarizing scientific evidence to address healthcare related issues such as disease treatment, prevention, diagnosis, and risk factor assessment (AHRQ, 2002). The steps involved in conducting a systematic review include: (1) development of a focused study question (2) identification of all potentially relevant primary research using a structured search strategy (3) screening of abstracts for relevance to the study question (4) quality assessment of the relevant literature (5) extraction of data from articles of sufficient quality and (6) synthesis of data from those studies using qualitative or quantitative approaches (NHMRC, 1999; CRD, 2001; Glasziou et al., 2001; Cochrane, 2004).

A systematic review differs from a traditional review in several ways. The process of systematic review reduces bias in the selection of research studies by the comprehensiveness and reproducibility of the search strategy and the transparent selection of articles included in review. Systematic reviews assess the methodological quality of the included studies (i.e. how well the study was designed, conducted and analyzed) and evaluate the overall strength of that body of evidence. In systematic reviews, emphasis is placed on the results from studies of higher quality rather than from lower quality. This additional analytic step does not typically occur during the course of narrative reviews. A summary table of the differences between systematic and narrative reviews is presented in Appendix 1.

An additional strength of systematic reviews is that by identifying all relevant and methodologically sound data, they improve the ability to synthesize the results of multiple studies and thereby increase power. Similar results observed across a wide variety of study designs and study settings provide evidence of robustness and transferability of those results to other settings. If the studies are inconsistent between settings, then the sources of variation can be examined (NHMRC, 1999; Glasziou et al., 2001).

Formats for conducting systematic reviews have been designed primarily for use by those who want to make more informed decisions in clinical practice, human healthcare research and public health policy. As such, numerous groups and organizations are involved in developing methodologies and conducting systematic reviews in the human health field. Examples of groups involved in the methodology and conduct of systematic reviews include:

1) The **Cochrane Collaboration** review format is an internationally recognized format for systematic reviews (Available at: http://www.update-software.com/ccweb/cochrane/hbook.htm).
2) The **Agency for Healthcare Research and Quality (AHRQ)**, previously the Agency for Health Care Policy and Research (AHCPR), provides research support and policy guidance in health services research and systematic reviews (Available at: http://www.ahrq.gov). In this role, the AHRQ places particular emphasis on the quality of care, clinical practice guidelines and evidence-based practice. This includes assessing the quality of published evidence through methods or systems to rate the strength of the scientific evidence for the underlying healthcare practice, recommendations in the research literature and health technologies.
3) The **Public Health Research, Education and Development (PHRED)** program conducts clinically relevant research into public health, health promotion and primary care and also fosters evidence-based practice and policy making (Available at: http://www.phred-redsp.on.ca/).
4) The **NHS Centre of Reviews and Dissemination (CRD)** produced an original document that provides the framework for carrying out systematic reviews of effectiveness (Available at: http://www.york.ac.uk/inst/crd/srinfo.htm). This document provides additional guidance on the evaluation of research relating to diagnostic tests, aetiology and risk factor studies, qualitative research and health economics.

The resources provided by these groups were instrumental in the preparation of this manual. An additional resource used in the creation of this manual was Systematic Reviews in Health Care: A Practical Guide by Glasziou et al., (2001). This book provides a clear and structured approach to systematic reviews. A unique feature of this book is the simple and clear descriptions of the various methods needed for different types of healthcare questions including the frequency of disease, prognosis, diagnosis, risk and management.

Despite the common use of systematic reviews in human heath related fields, formal systematic reviews have rarely been used in agriculture and agri-food safety. One example is provided by Australian researchers, who applied this approach to evaluate the evidence for and against the use of antimicrobials in animals as a contributor to the emergence of clinically significant disease in humans (Ferguson et al., 1998).

The protocols for developing systematic reviews in human health are well developed. However, systematic review protocols developed for use in human health studies may not be directly applicable to evaluate agriculture or agri-food safety issues at the farm level. For instance, many on-farm intervention studies use an observational study design compared to randomized controlled trials, a study design frequently used for systematic reviews in the human healthcare field. Challenge studies conducted in veterinary research are generally not a design considered in human health reviews. In the human healthcare field, challenge trials are conducted in non-human species whereas, in veterinary research, this type of trial is conducted in the species of interest. Another factor to consider is that livestock populations are grouped in a very different way compared to human populations. This means that statistical issues related to non-independence of study subjects within groups are of paramount importance in many agri-food research studies. Therefore, it is necessary to modify the existing protocols used for systematic reviews in the human health field for use in systematically evaluating agri-food research.

The use of systematic reviews will allow researchers to synthesize the current body of knowledge on targeted food safety issues and lend increased credibility to findings in the field. The findings of such independent reviews can offer valuable information on the best interventions and can provide data as input into risk assessment models. Systematic reviews also can highlight areas where there is insufficient evidence of the efficacy of interventions or where there are common methodological flaws in the available research and thereby provide direction and impetus for future basic and applied research in a specific food safety area.

The purpose of this manual is to provide guidelines and recommendations for conducting systematic reviews in the agri-food safety area. While systematic reviews may be conducted to address a broad range of research questions, such as interventions, disease incidence / prevalence estimates, diagnostic test comparisons, program evaluations, and questions of associations, the emphasis of this manual is the evaluation of intervention research. The manual also emphasises the pre-harvest (on-farm) food safety component of agri-food public health. The manual is structured to work through the steps of conducting a systematic review, namely:

1) Development of a focused study question
2) Identification of all potentially relevant primary research
3) Screening for relevance
4) Quality assessment
5) Data extraction
6) Data synthesis

Example forms for each step are provided, and additional details are included as appendices. To illustrate the concepts, a working example of a systematic review, "The use of probiotics to reduce *E. coli* O157 in the feces of beef and dairy cattle" is used throughout. Our research group has used the web-based systematic review software

"Electronic Systematic Review" (ERS) (www.trialstat.com, O'Blenis and Garritty, 2004) to manage systematic review projects. Comments on the use of this software will be included in this manual.

1. Development of a focused study question

KEY POINTS

- **The question should be clearly defined *a priori***
- **The question should be structured in terms of population(s), intervention(s), and outcome(s)**
- **The question should be sufficiently broad to allow examination of variation in the study factor and across sample populations**
- **The system level or sector of agriculture within which the review will be conducted should be specified**
- **If sufficient literature is available, the review may be structured to include only study designs that provide a higher level of evidence**

Defining the question for a systematic review is critical because all other aspects of the review flow directly from that question (CRD, 2001). Specifically, the question guides the search strategy for identifying potentially relevant studies, for determining relevance of the studies identified by the search strategy, for critically appraising the studies, and for analyzing variation among results (Cochrane, 2004).

1.1. Formulating a relevant question

When framing precise questions the important facets to be considered are:

- Population(s)
- Intervention(s)
- Outcome(s), and
- System level or agricultural sector

Differences in these characteristics will have an impact on the effectiveness of the interventions being reviewed. Therefore, it is important to specify *a priori* the important population characteristics, the intervention strategies acceptable for inclusion in the review, and the most clinically relevant outcome(s) to measure the effect of the intervention of interest (CRD, 2001).

1.1.1. Population

Population characteristics that vary between studies in agri-food public health research include species, age of animals, production system, and country. Defining the review question in terms of these population characteristics will aid the reader in determining the relevance of the results of the systematic review to the population to which they would infer the results.

1.1.2. Intervention

There are many types of interventions that may be the subject of a systematic review in agri-food public health such as:

- Therapy for a specific animal disease or pathogen
- Prevention of a specific animal disease or contamination of a product
- Management interventions to decrease exposure to one or more pathogens

The defining feature is that a specific activity is undertaken with the aim of improving or preventing adverse health outcomes.

1.1.3. Outcome

The review should address relevant and important outcomes that are meaningful to individuals making decisions about agri-food public health interventions. The initial literature searches will identify the types of outcomes used in primary studies to measure the response to an intervention (CDR, 2001). The outcomes can be:

- Qualitative
- Quantitative
- Economic

1.1.4. System level or agricultural sector

It is important to determine and consider the system level or sector of agriculture within which the systematic review will be conducted. Such sectors may include a specific livestock commodity group (animal, farm, or herd level), processing, retail, or consumer studies, or may include more than one level of the farm to fork continuum. Within a livestock commodity, the intervention may be targeted to a specific production group (e.g. pre-weaned animals, finishing animals, cull animals).

1.2. Focusing the question

The overall objective of conducting the systematic review may be to evaluate interventions in general to reduce a specific pathogen. In this case, the reviewer may want to specify multiple research questions and conduct a systematic review for each of these. For instance, if the overall objective is to identify interventions to reduce fecal shedding of *E. coli* O157 in beef and dairy cattle, a single question may not have sufficient focus for the review. Therefore, the objective should first be framed into the component parts. In this example, the issue could be divided into (1) factors that increase animal resistance (e.g., vaccination, bacteriophages, medicinal plants or other feed additives) and (2) management interventions to reduce exposure and/or transmission (e.g., cleaning of pens, chlorination of water). Specific questions can be developed within each of the component parts. For our working example of probiotic use to reduce *E. coli* O157 in cattle, a review question related to probiotic use would be one question within the 'animal resistance' component.

When framing our question we also need to consider how narrow or broad it should be. The question should be sufficiently broad to allow examination of variation in intervention effects across a range of relevant populations. Consider the following question related to probiotic use in cattle:

What is the effect of the use of *L. acidophilus* on *E. coli* O157 in the feces of post-weaned ruminants?

This question addresses all of the key components of a suitable review question. The population is specified as post-weaned dairy and beef cattle, the intervention is defined as the use of *L. acidophilus*, the outcome is fecal shedding of *E. coli* O157 (although the question does not clearly state whether this is prevalence, incidence or bacterial load), and the sector, on-farm, is implied. However, this question might be too specific as an initial review question because there is not much information about the use of specific probiotics to reduce *E. coli* O157. A better question in this case may be:

What is the effect of the use of probiotics on *E. coli* O157 in the feces of post-weaned ruminants?

If the terminology of the question is ambiguous or if all of the members of the research team are not familiar with the exact meaning of the words, the question may include further clarifications. For example:

What is the effect of using probiotics on *E. coli* O157 in the feces of post-weaned beef and dairy cattle?
where "probiotics" include: commensal (harmless or beneficial) bacteria that are administered to reduce pathogenic bacteria in the gut. These include *Lactobacillus*, *Streptococcus, Enterococcus, Bifidobacterium and Saccharomyces*. The terms 'competitive exclusion' and 'strain mixture' should be included in this study, and
where *E. coli* O157 will include the terms *E. coli* O157:H7, enterohaemorrhagic *E. coli* (EHEC), verotoxigenic *E. coli* (VTEC), and Shiga-like toxin *E. coli* (STEC), and
where ruminants include either dairy or beef cattle, goats or sheep.

1.3. Study designs and levels of evidence

Research questions can be addressed using a number of study designs. The different study designs have advantages and disadvantages but, in the context of using data from different study designs, it is important to consider the level of evidence that each design provides.

Randomized controlled trials represent an experimental design where individuals (or study units) are randomly allocated to treatment groups. The disease challenge is natural and the trials are conducted in a natural setting.

Observational studies relate individual characteristics, personal behaviours, environmental conditions and treatments as 'exposures' that may modify the risk of disease. In contrast to randomized trials, observational studies relate to naturally occurring exposures to disease and a natural disease challenge. Observational studies are

differentiated by the method of selecting study populations and may be (1) cohort, (2) case-control or (3) cross-sectional. Cohort studies relate exposure to the subsequent onset of disease and compare the incidence of disease among exposed and unexposed populations. Case-control studies compare the exposure histories of a group of cases to those of controls (disease free). Cross-sectional studies determine the exposure and outcome status of subjects at a particular time. The point of time may range from an instant ('time of sampling') to longer periods (such as "during the past year"- longitudinal cross sectional studies) although all are treated as static, point-in-time events.

In the field of aetiology, cohort studies provide stronger evidence for testing hypotheses than case-control studies. Cross-sectional studies are used to generate hypotheses (as opposed to testing hypotheses). Therefore, they are generally of less importance in a systematic review, although in certain reviews they may be useful.

One study design commonly used in pre-harvest food safety research is the challenge or inoculation study. Challenge studies have an experimental design that features deliberate inoculation with the disease agent of interest in animals randomly assigned to exposure groups. These studies provide evidence related to the efficacy of an intervention under controlled conditions. However, they may not be representative of the efficacy of that intervention under commercial settings and with a natural disease exposure.

Descriptive studies are generally not included in systematic reviews, as they are used to provide baseline or preliminary data, and not appropriate for hypothesis testing.

Table 1 summarizes the level of evidence obtained using different study designs to answer specific scientific questions. A grading system is included, with level I representing the highest level of evidence.

Table 1. **Grading the level of evidence obtained using different study methodologies to answer specific scientific questions.**

Level of evidence	Study design	Study type
I	Systematic Reviews	Structured review
II	Randomized clinical trials	Experimental
III	Cohort studies	Observational
IV	Challenge trials	Experimental
V	Case-control studies	Observational
V	Cross-sectional studies	Observational
VI	Descriptive studies, case reports / case series, opinion of respected authorities, reports of expert committees.	Descriptive

Adapted from Ferguson *et al.* (1998) and Glasziou *et al.* (2001)

When conducting systematic reviews of human health treatment interventions, there are often a sufficient number of randomized controlled trials (RCTs) to allow the systematic review to be restricted to this type of study design. However, in the literature related to on-farm intervention strategies, this may not be the case. Often, the research evidence to address on-farm interventions comes primarily from observational studies and challenge trials although, in some instances, RCTs provide critical tests of causal hypotheses. Thus, it may be necessary to include studies from different "evidence levels" in an agri-food public health systematic reviews. However, it is difficult to combine data from studies conducted at different levels of evidence. Therefore, if there are an insufficient number of level II studies available, an alternative approach is to include multiple study types in the review and summarize the results within each evidence level.

2. Identification of all potentially relevant primary research

KEY POINTS

- **The objective is to generate a complete list of all primary research (published and unpublished) that could potentially answer the research question.**
- **Effective combinations of search terms are constructed using the key components of the review questions (population, intervention, outcome, and agricultural sector).**
- **Identification of relevant literature is performed by initially searching electronic databases, and consequently searching reference lists and obtaining data from unpublished studies when possible.**

The purpose of the search strategy is to generate a comprehensive and complete list of all primary research (both published and unpublished) that could contribute to answering the question posed in the review. The identification of studies by a thorough and unbiased search strategy is crucial. This is because the validity of the review findings is directly related to the comprehensiveness of the search used to capture potentially relevant studies and the reproducibility of the search protocol.

The choice of a sensitive versus a specific search strategy will depend on the purpose of the systematic review. A fully comprehensive review requires high sensitivity literature identification as a starting point. Irrelevant studies will be removed from the systematic review process at the relevance screening stage described in section 3 of this manual. However, if the volume of available abstracts is very high, then a more restrictive search strategy may be appropriate (i.e. the inclusion only of study designs representing higher levels of evidence, as described in section 1).

Search strategy may be developed as an iterative process. Initially, a trial search is performed. The results of this search are discussed within the review team and also may involve content experts to ensure that all potentially relevant search terms are included. Successful search strategy design involves knowledge of databases, indexing and database text structures. Hence, successful search strategies typically involve experienced information specialists (CRD, 2001).

2.1. Creating search terms

Constructing an effective combination of search terms for searching electronic databases requires a structured approach. The initial approach involves searching different databases for combinations of the intervention and outcome of interest in certain populations. If the resulting set is too large, a methodological filter may be used (i.e. the search may be restricted to study designs representing only evidence level II studies). The following steps are used to create the search terms:

- Break down the study question into components (population, intervention, outcome, agricultural sector)
- Identify search terms in each component that best capture the subject and identify which terms may be a subset of other more important terms. This will help to focus the search
- Combine search terms within each component using "**OR**" (to ensure that all records with at least one of the specified terms are identified) and combine components using '**AND**' (to ensure that all the components must appear in the record)

Examples of component combinations include:

1. Population **AND** intervention
2. Intervention **AND** outcome
3. Population **AND** outcome
4. Population **AND** intervention **AND** outcome

- '**NOT'** can also be used to exclude records from a search. For example, 'probiotic' **NOT** 'biophage' will retrieve all records that contain the term 'probiotic' but not those that also contain the word 'biophage'. **NOT** should be use with caution because it may have a larger exclusion effect than anticipated (as it may exclude records of interest that coincidentally discuss both terms)

In general, fewer components will result in more articles being identified but will increase the number of non-relevant publications.

For the example of the use of probiotics for the reduction of *E. coli* O157 shedding in cattle it was found that:

(Lactobac*) AND (cattle) AND (*Escherichia coli* O157)

was very specific but much relevant research was excluded.

Whereas the search strategy:

(Lactobac* OR yeast OR dietary supplementation) AND (ruminant OR beef OR dairy) AND (*Escherichia coli* OR enteric OR feces*)

will result in a more sensitive search process.

NB: using a partial word followed by * will identify all words containing the partial text. For example, diet* will identify diet, dietary, and diets.

The aim of creating a search term list is to identify combinations of search terms within each component that will maximize identification of potentially relevant articles. One way to begin identifying vocabulary terms for a review is to retrieve a few subject-related abstracts and note commonly used text words and keywords that indexers have applied to the articles. Appendix 2, Table 1, illustrates the number of abstracts identified using population, intervention, and outcome terms potentially relevant to the probiotic example. Adding more than 1 search term within a component, and linking the multiple terms with "OR", will increase the number of abstracts identified although, due to the use of multiple terms within many abstracts, the resulting numbers are not cumulative (see Appendix 2, Table 2). Appendix 3 includes several tables with some of the general and specific terminology commonly used for different populations, interventions and outcomes related to agri-food public health. Appendix 4 shows the number of abstracts identified using various population search terms in 3 electronic databases for cattle, swine, and poultry.

In order to identify relevant search terms, an exclusion of individual terms may be performed where all the search terms within a component are initially included. The exclusion of terms is performed in a backwards fashion by excluding each term one-at-a-time. If the number of papers obtained is reduced after excluding a specific term, then that term should be included in the final search term list. For our probiotic example, including both 'EHEC' and 'shiga-like toxin" increased the number of abstracts, but including 'VTEC' with the other terms did not increase the number of abstracts identified.

Maximizing the sensitivity of the search means that many of the articles identified may not be relevant to the review question. The amount of irrelevant material can be reduced substantially by using a methodological filter. For example, this may be done by focusing on types of studies that are most likely to yield sound data relevant to a given population. Many human medical systematic reviews are restricted to randomized control trials. However for pre-harvest food safety reviews, randomized control trials may not be numerous and hence, including data from multiple study types may be necessary. A free MEDLINE facility that uses methodological filters, developed by Haynes et al., (1994) is available from the National Library of Medicine (available at: www.nlm.nih.gov/) in the section on PubMed-Clinical Queries.

Additional considerations:

- When defining search terms, one should consider the use of the singular and/ or plural versions of the search terms. In our probiotic example, searches using the word 'probiotic' identified 2055 papers while the word 'probiotics' identified 1897 papers. However when both terms were combined in one search, the number of papers obtained was 2251. This indicates that each term individually captured some papers that the other term did not. In addition, different spellings and acronyms for the search terms should be considered (e.g. faecal and fecal, VTEC and verotoxigenic *Escherichia coli,* EHEC and entero-haemorrhagic *Escherichia coli*).

- Based on the history of research of the specific intervention and/or outcome investigated, it is possible to limit searches to specific years. It is known for example, that *Escherichia coli* O157 emerged as an important human pathogen in the early 1980's. Moreover, the use of probiotics in cattle to prevent the shedding of bacteria is a fairly novel procedure. Because of this, the literature search for the probiotic example was restricted to papers published after 1980.

- It is important to decide which languages are to be included in the search. A language filter may be used to exclude non-English references. One could also choose to have no language screen and to later decide which articles to translate when the volume of literature in different languages is determined. Restriction of language can introduce bias and decrease the precision of the systematic review. Trials performed in some countries may be more likely to have positive results, reflecting publication bias based on geography. Therefore whenever possible, all suitable reports should be included regardless of language (CRD, 2001). It may be worth considering the use of search terms in languages other than English in some databases.

Table 2 contains relevant terms used in the review: "The use of probiotics for the reduction of *E. coli* O157 in the feces of post-weaned beef and dairy cattle".

Table 2. Relevant within and between population, intervention and outcome terms for the study "Effect of probiotics on the reduction of *E. coli* O157 in the feces of post-weaned beef and dairy cattle."

Population	Intervention	Outcome
Ruminant	Probiotic	*Escherichia coli*
Ruminants	Probiotics	*Escherichia coli* O157
Bovine	Lactobac*	O 157
Cattle	Bifodobac*	"Bacteria load" or "bacterial load"
Cow	Propionibac*	"Bacteria level" or "Bacterial level"
Cows	Saccharomyces	"Bacteria log"
Steer	Competitive exclusion	"Bacteria counts"
Calves	Fermentation	Faeces
Beef	Strain mixture	Manure
Farm	Dietary supplementation	Gastrointestinal
Herd	Yeast	Fecal
Dairy	Lactic acid	Feces
	Bacteriocin	Coliform
	Lactic acid bacteria	*Enteric*
		EHEC
		Shiga-like toxin

The purpose of our question was to determine the effectiveness of probiotics for reducing *E. coli* O157 shedding at the farm level. A significant amount of probiotic research has focused on improving the safety of meat products after slaughter or to prevent the contamination of milk products. Therefore, the term **'NOT'** was added to terms such as milk and cheese to exclude them from our search strategy.

2.2. Electronic databases

The most efficient means of identifying potentially relevant studies is through the use of health-related or agricultural electronic bibliographic databases. These databases can be searched for specific words in the title and abstract or for standardized subject-related indexing terms assigned to the records of the database. For example the term RANDOMIZED-CONTROLLED-TRIAL was introduced in MEDLINE in 1991 and allows the user to search for articles that describe this study design. Most electronic bibliographic databases include abstracts for the majority of recent records and many include links to electronic copies of full articles (Cochrane, 2004).

There are many potentially useful databases and guides to databases that can be consulted in health care. The electronic databases generally considered as the richest sources of references in medicine, agriculture and agri-food public health are MEDLINE, AGRICOLA and CAB International, EMBASE, FSTA, INGENTA, and Biological Abstracts (BIOSIS).

1) **MEDLINE** indexes approximately 4600 journals from the United States and 70 other countries. It includes abstracts and, for the majority of veterinary medicine and human medicine publications, it includes full text articles. PubMed is a free, online MEDLINE database (Available at: http://www.ncbi.nlm.nih.gov).

2) The AGRICultural OnLine Access (**AGRICOLA**) database is a bibliographic database of citations to the agricultural literature created by the National Agricultural Library (NAL). AGRICOLA includes publications and resources related to agriculture and allied disciplines including animal and veterinary sciences, plant sciences, forestry, aquaculture and fisheries, farming and farming systems, agricultural economics, extension and education, food and human nutrition, and earth and environmental sciences. Although the AGRICOLA database does not always contain full publications, thousands of AGRICOLA records are linked to full-text online documents. AGRICOLA is searchable online at http://agricola.nal.usda.gov and may be accessed on a fee basis through several commercial vendors. Alternatively, AGRICOLA files may also be leased from the National Technical Information Service (NTIS).

3) **CAB International** is a comprehensive file of agricultural information containing all records from the more than 50 journals published by CAB International (CABI). Of particular note are sections in the database covering literature in the fields of veterinary medicine, agriculture (AGORA available at:

http://www.aginternetwork.org/en/journals.php) and human nutrition. In addition to the over 14,000 serial journals in more than 50 languages, CAB international also contains books, reports and other publications. CAB abstracts are available through CAB Direct at: http://www.cabdirect.org/ from CABI Publishing or online through a variety of service providers.

4) **EMBASE** is a comprehensive index of international literature on human medicine and related disciplines. Approximately 500,000 records are added annually, in recent years over 80% of which contain abstracts. EMBASE provides access to periodical articles from more than 4,600 primary journals from approximately 70 countries. An additional 350 journals are screened for drug articles. EMBASE consists of two files: File 73 contains records from January 1974 to the present; File 72 contains records from 1993 to the present. ONTAP® EMBASE. File 272, is available for ONline Training and Practice (Available at: http://www. embase.com/search, requires user name and password)

5) **FSTA** (Food Science & Technology Abstracts) provides access to international literature on every aspect of food science, food products and food packaging. It covers 1800 scientific journals as well as patents, books, conference proceedings, reports, pamphlets, and legislation. All the abstracts included are in English and are prepared by specialists, from works originally published in more than 40 languages. In electronic form, FSTA covers the literature produced from 1990 to the present. The University of Guelph Library holds a complete set of Food Science & Technology abstracts in paper format from 1969 until 1998 (Available at: http://web5.silverplatter.com.cerberus.lib.uoguelph.ca/webspirs/start.ws?customer=tug&databases=(FSTA).

6) **INGENTA** Library Gateway is a searchable database of more than 11 million citations from over 25,000 journals. This database includes journal articles and current research (Available at: http://www.gateway.ingenta.com.cerberus.lib.uoguelph.ca/uoguelph) (U of Guelph) or (http://www.ingenta.com).

7) **BIOSIS** (Biological Abstracts) provides access to nearly 6,000 international journals related to agriculture, biochemistry, biotechnology, ecology, immunology, agriculture, biochemistry, biotechnology, ecology, immunology, microbiology, neuroscience, pharmacology, public health and toxicology (Available at: http://isi02.isiknowledge.com/portal.cgi).

8) **ISI Web of Knowledge** is an integrated web-based platform providing high-quality content and the tools to access, analyze, and manage research information. These tools include cross-product searching, links to full text, citation alerts, table of contents alerts, personal journal lists and personal bibliographic management (Available at: http://isi02.isiknowledge.com/portal.cgi).

As the number of databases used in the electronic searching increases, the number of abstracts identified will also increase, albeit with diminishing returns. However, the aim of the electronic database search is to identify *all* of the potentially relevant literature and each of these databases includes some unique journal sources. Therefore, it is recommended that multiple databases be used.

Different electronic databases have different formats for entering the search terms. Appendix 5 uses the probiotic example to illustrate how search terms are entered into the different electronic databases. For example, when combined terms are considered as a phrase (such as 'bacterial load'), it is necessary to use quotations "" in the PubMed database, while in Agricola and CABI it is not. A library resource person can be very helpful at this stage to assist with formatting search terms for different databases.

Systematic review reports must document the search process. For electronic database searches, provide a copy of the search strategy for each database used including:

- Name of database searched
- Name of host/system used to access the database, for example through PubMed, Silverplatter or ERL Webs SPIRS.
- Date of search
- Years covered by the search
- Language(s) included in the search
- Complete search strategy used including ALL search terms
- Number of articles retrieved using each specific search term in each database and the total of articles retrieved within each of the components.

Example of search strategy documentation:

> MEDLINE, PubMed, search December 21, 2004, period 1980 to 2004 search terms: (probotic or probiotics or "lactic acid bacteria" or "direct fed microbial" or "strain mixture" or "live bacteria supplement" or "competitive exclusion" or Lactob* or Bificobat* or Propionibact* or bacteriocin) and (bovine or cattle or beef or cow or cows or steer or steers or heifer or calf or claves or ruminant) and (O157 or O 157 or "*E. coli* log counts" or "*E. coli* CFU" or EHEC or VTEC or STEC or enterohaemorragic or verotoxigenic or shiga-like toxin or faecal or fecal or coliform or enteric). Number of abstracts: 159

NOTE: In our probiotic example, we were aware of several observational studies that considered probiotic use as one of a large number of potential risk factors. These studies were not identified by our search strategy. Thus, an additional search of observation risk factor studies for intervention strategies for *E. coli* O157, without regard specifically to probiotic use, was conducted using the following terms: ("risk factor*" or management) and (cattle) and (coli*). The search was performed using the same protocol and databases as for the original search strategy described for this example review.

An efficient way to manage this extensive volume of literature is to download the journal citations identified by the search, and their abstracts, into a commercial reference

manager (e.g. Procite, Reference Manager). Most reference management software programs have built-in programs to identify and remove duplicate records (NB: in some cases, only identical records will be identified. We have found that records from the same study with different citation formats will not always be captured as duplicates, nor will multiple publications using the same data – such as publication of preliminary versus final results). We are downloading citations and abstracts from Reference Manager into a commercial web-based software, ESR (Electronic Systematic Reviews), for project management of the systematic review. This program provides a format for data management of all stages in the review and has its own de-duplication program.

Electronic database searches may not identify all of the relevant literature. Therefore, additional search strategies are recommended. These strategies include hand searching of journals and conference proceedings, checking reference lists and identifying un-published data. These approaches are further elaborated in the following sections.

2.3. Hand-searching

Hand-searching involves an examination of the contents of a journal issue or conference proceedings to identify all eligible reports whether they appear in articles, abstracts, proceeding abstracts, news columns, editorials, letters or other text. It is important to conduct this type of search because not all relevant literature will be included within electronic databases or, if included, may not be indexed with terms that allow identification (Cochrane, 2004). Potential journals that may not be indexed in electronic databases should be identified *a priori*. Hand-searching should be documented using the full title of the journal and the first and latest years searched. Any issues not searched because of missing journal issues should be recorded. For example:

> Journal of J Vet Med B Infect Dis Vet Public Health 1970-2003
> Missing issues 2 (1978), 3 (1973)

Conferences that may contain relevant information should be identified *a priori*. The indexes of these conferences should then be hand-searched for relevant abstracts. In many cases the abstracts may contain sufficient detail to enable relevance screening (see section 3). However, for abstracts that are deemed to be relevant it may be necessary to contact the authors for additional details of the study in order to allow quality assessment to be performed. Details of conference proceedings searches should be documented as follows:

1) Proceedings with a title in addition to the conference name:
   Author, Title. (2003). Proc. 5th International Symposium on Shiga Toxin-Producing *Escherichia coli* Infections, Edinburgh, U.K.

2) Proceedings also published as part of a journal:
   Author, Title. (2003). J. Anim Sci, 81 (E Suppl. 2).

2.4. Checking reference lists

The reference lists of selected relevant articles should be examined to ensure that the search strategy has identified all potentially relevant studies and thereby provides a means of validating the electronic database search. Additionally, the process of following up references from one article is another means of identifying studies for possible inclusion in a review (Cochrane, 2004). Finding the 10 most recent articles published on the topic and checking the references back against articles identified by other methods and finding the 10 oldest papers on the topic and using a citation search from that date forward is an objective way of checking reference lists.

2.5. Finding unpublished studies

Some completed studies are never published, and there is a lag time from completion of a study to publication. The association between significant results and publication has been well documented and it is therefore important to search for unpublished studies to minimize bias in the systematic review (Dickersin, 1997). Unfortunately, it is difficult to obtain information from research that has been completed but never published.

Informal channels of communication, such as contacting researchers in the area of the review, can sometimes be the only means of identifying unpublished data (NHMRC, 1999; Cochrane, 2004). To maintain the reproducibility of the search strategy, specific criteria for identifying and contacting the content experts should be established *a priori* and documented in the final report with a list of all content experts contacted by the review group.

Identifying ongoing studies also is important. Databases that may contain information on this source of information are:

- TrialsCentral (available at: www.trialscentral.org)
- Current Controlled Trials (available at: www.controlled-trials.com)
  NB: these two groups provide central access to ongoing human trials.
- The inventory of Canadian Agri-Food Research (ICAR) (Available at: http://ontarioandnunavut/hpcb/ontregions.nsf/Introduction?OpenForm).
- The United States Department of Agriculture (USDA) (Available at: http//www.ars.usda.gove/research/projects.htm)
- TEKTRAN: contains published or soon-to-be-published articles of recent research results from the Agricultural Research Service (ARS), the U.S. Department of Agriculture's chief scientific research agency. TEKTRAN is updated regularly, adding summaries of new articles accepted for publication and removing summaries after five years. Some summaries are excluded until the article appears in publication. When the articles are published, the publication dates are added to TEKTRAN. This database also contains scientists' contact information and additional links to their respective research projects and other publication submissions (Available at: http://www.nalusda.gov//ttic/tektran/tektran.html).

- Current Research Information System CRIS is the U.S. Department of Agriculture's (USDA) documentation and reporting system for ongoing and recently completed research projects in agriculture, food and nutrition, and forestry. Projects are conducted or sponsored by USDA research agencies, state agricultural experiment stations, the state land-grant university system, other cooperating state institutions, and participants in a number of USDA-administered grant programs (Available at: http://cris.csrees.usda.gov/).

These resources can be an important source of current and ongoing research, but tend to have limited search capacities and thus may require a modification of the search strategy.

To document unpublished studies, provide a brief summary of those databases searched and efforts made to contact investigators. For example:

> United States Department of Agriculture (USDA) Research Projects database (http://www.ars.usda.gove/research/projects.htm)
> Research Project: Prevalence of *E. coli* O157:H7 on hides and in feces of feedlot cattle over time. Project team: Koohmaraie, Mohammad (April, 2004-December, 2005).
> Searched: January 19, 2005.

Another way of finding unpublished literature and ongoing research is by searching databases that contain information on thesis and dissertations in progress. Some databases recommended for this purposes are:

- National Library of Canada. Thesis Canada Portal (http://www.collectionscanada.ca/thesescanada/index-e.html).
- Canadian, U.S. & European (1861 to present), view abstracts and 24 pages of the thesis (1997 to present) (http://wwwlib.umi.com/dissertations/search).
- The universal index of doctoral dissertations in progress (http://www.phddata.org/).
- The Current research@ ProQuest Digital Dissertations. UMI's Dissertation Abstracts database (http://wwwlib.umi.com/cresearch/gateway/main). It can be browsed by name or location of institution (university).

One should provide documentation of any other sources searched (eg. Bibliographies and Internet websites). Documentation should include:

- Details of date searched,
- Search terms used,
- The URL (if relevant),
- Any specific features of the resources that might influence the search process (such as 'only the titles were searchable' or 'word sets could not be combined').

When conducting searches of the unpublished literature, the information obtained may consist only of abstracts or outlines of currently funded research projects and therefore may have insufficient detail for relevance screening or for quality assessment, should the information be deemed relevant. In this instance, it may be necessary to contact the investigators to request additional detail. If sufficient detail is available to allow relevance screening, contact may be made only for those references that pass screening. A generic investigator contact letter is provided as Appendix 6.

3. Screening abstracts for relevance to the study question

KEY POINTS

- **The relevance screening tool is used to quickly determine if an article may be relevant to answer the research question.**
- **Reviewers answer a series of questions based on information available in the abstracts.**
- **Abstracts should be assessed independently by at least 2 reviewers and agreement between reviewers should be evaluated to ensure reproducibility.**

The aim of the search strategies is to identify all potentially relevant research that may address the systematic review question (i.e. high sensitivity, but potentially low specificity). The purpose of abstract screening for relevance is to identify those articles that specifically may help to address the review questions. At this step, it is essential that decisions about inclusion or exclusion of studies are made according to predetermined written criteria in order to prevent any bias in decision-making. If an abstract passes the relevance screening step, a copy of the full article is obtained and the reference moves on to the quality assessment step (see section 4). Abstracts that do not pass relevance screening are not considered further in the systematic review.

3.1. Developing the relevance screening tool

The relevance screening tool generally consists of a short series of questions which are designed to quickly determine whether or not the article belongs in the review. As systematic reviews generally only consider primary research, the first question may be used to eliminate review articles (other than previous systematic reviews). The remaining questions should be specific and defined based on the population, the intervention, the outcomes and the study designs of interest. If the review is considering only one specific question, then the screening questions should pertain to each of the key component of the review question. In the section on 'Systematic Review Forms', Form 1 includes a generic Abstract Relevance Screening tool used for a single question review. The relevance screening tool for the probiotic example is provided as Form 2. With this type of review, the screening tool questions are usually structured such that studies fulfilling all of the inclusion criteria (i.e. answer all of the questions affirmatively) pass the relevance screening step.

In some instances, the systematic review may address several related questions. An example of this is a review currently being conducted by our group on the evidence for an association between Johne's disease in cattle and Crohn's disease in humans. In this example, there are 3 specific questions related to the presence of the putative pathogen (*Mycobacterium avium subsp. paratuberculosis*) in human cases. Each of these questions essentially becomes it's own systematic review. In this type of review, the relevance screening tool has one question related to whether the study is primary research and one

additional question which incorporates each of the specific questions in the review (see "Systematic review forms", Form 3). To pass relevance screening, the reviewers must identify the abstract as primary research and answer in the affirmative for at least one component within the other question.

3.2. Use of the relevance screening tool

The relevance screening tool is used to determine if an article may be appropriate for the systematic review. This involves screening of only the titles and abstracts of the papers. The relevance screening tool is applied to each of the citations generated from the search strategy.

Abstracts should be independently assessed by at least 2 reviewers to increase the reliability of the inclusion and exclusion process. If both reviewers determine that the abstract does not meet inclusion criteria, it can be rejected. If the two reviewers disagree, the conflict should be resolved by consensus or, if this is not possible, by a third reviewer. Disagreements and their resolution should be documented. When a reference is passed as potentially relevant, then a full-text copy of the reference is obtained and the article moves on to the quality assessment stage. The references that do not meet the screening requirements are eliminated from the review. The references for which screening questions cannot be answered due to lack of information should remain in the study and efforts should be made to obtain further details of the study.

The reproducibility of this process should be tested in the initial stages of the review (we evaluated agreement after all of the reviewers had screened the same 50 abstracts). If reproducibility is shown to be poor, rewording of the questions with more explicit criteria should be undertaken. The agreement between reviewers can be statistically evaluated using Cohen's Kappa.

If there are a large number of abstracts identified by the search strategies, the two reviewers may screen abstracts in parallel (i.e. the abstract passes relevance screening and moves on to the quality assessment stage as soon as one reviewer passes the abstract). This can greatly reduce the time necessary for the relevance screening stage. However, duplicate relevance screening should still be done initially to allow validation of the relevance screening questions, as described above.

The ESR software allows not only the storage and distribution of references and full documents but also the design of electronic review forms, the assignment of reviewers to specific citation subsets or forms, the capture and reporting of reviewers input, automated inclusion and exclusion of references, and evaluation of reviewer agreement.

4. Quality assessment of the relevant literature

KEY POINTS

- **Quality assessment is a second stricter level of screening where reviewers have access to full copies of the references.**
- **The quality of each article is evaluated using a standardized procedure (checklist) for each type of study design.**
- **The quality assessment tool is composed of generic and individual components of study methodology that have a potential relation to bias (selection, confounding, and losses to follow-up) and validity of the studies.**
- **Components of the quality assessment that are essential to define minimum acceptable quality are identified, and articles that do not meet these criteria are excluded from the review.**

Interpretation of study results depends on the design, conduct, and analyses (internal validity) and the population, interventions, and outcomes (external validity). Combining results or effect measures of interest in a review may be biased if studies of varying quality are summarized together. Therefore, it is necessary to determine a minimum quality threshold for inclusion of research findings in the review. The purpose of the quality assessment step of a systematic review is to exclude studies whose quality is too low to provide meaningful data to address the review question.

Final inclusion/exclusion decisions should be made after retrieving the full text versions of all potentially relevant citations. A critical appraisal of the full article for each of the relevant studies is therefore needed, so that studies that are of appropriate quality can be selected. To avoid a selection that is biased by preconceived ideas, it is important to use a systematic and standardized approach to the appraisal of these studies.

4.1 Developing the quality assessment tool

A systematic review should base its conclusions on the highest-quality evidence available. For this purposes it is important to develop a valid and standardized procedure to select from the large pool of studies potentially identified for the review. Quality assessment instruments are based on individual aspects or components of study design (objectives, population, intervention, outcomes assessment, withdrawals and data analysis). These items are then assembled into one or more checklists, which are used to systematically evaluate each study.

The different study designs share many criteria for validity, although some elements essential to validity are specific to a particular type of study design. Therefore, when developing checklists, there will be elements common to all of the study designs and those specific to one or more designs. Appendix 7 shows the domains and elements of importance to quality assessment for randomized controlled trials and for observational

study designs. The Research Triangle Institute-University of North Carolina Evidence-Based Practice Centre prepared this summary for the Agency of Healthcare Research and Quality after summarizing more than 100 sources of information on systems for assessing study quality and strength of evidence for systematic reviews.

These checklists are used to ensure that studies have adequately addressed all of the components necessary to evaluate quality. The quality issues are usually grouped under the headings of generic and specific items. Generic quality items included in the checklists developed by our group are:

1. Objectives and study population
2. Intervention
3. Outcome assessment
4. Withdrawals and loss to follow-up
5. Data analysis and control for confounders
6. Conclusions

Individual aspects of study methodology are incorporated within the generic components for all study designs, for example information on precision of the study, external validity, randomization, comparison groups, blinding of intervention status, loss to follow up, assessment of relevant outcomes, and statistical analysis. It is important to consider these aspects because they have a potential relation to bias in estimation of effects**.**

The section 'Systematic review forms' (Forms 4 A, B, C, D) includes the quality assessment checklists developed by our group to evaluate the quality of randomized controlled trials (A), challenge trials (B) cohort observational studies (C), and case-control and cross-sectional observational studies (D) for answering questions related to on-farm interventions on agri-food public health topics. These quality assessment forms will generally be useful for all reviews of agri-food public health interventions. However, some questions, such as the one related to minimum laboratory standards for identifying specific pathogens and the one related to confounding, will need to be tailored to individual systematic reviews. Whenever possible, identical questions were used in each form to allow consistency of the quality assessment across study designs. An initial question pertaining to the suitability of the study objective for answering the review question is included as a check of the relevance screening. Although checklists have been provided for each of the major study designs, if sufficient information is available to address the review question, the systematic review may be restricted to study designs that provide higher levels of evidence. The weakest study design that may be included in the review should be clearly stated in the review protocol.

Individual criteria within each checklist can be evaluated in different ways such as "strong", "moderate" or "weak" or alternatively by stating that the criteria were "met", "not met" or "unclear" or simply by using "yes", "no" or "partial". In the checklists developed by our group, yes, partial and no were used to complete the evaluation. The criteria for each individual element to be considered as a yes, partial or no are specified in the forms.

An explicit set of criteria must be developed to determine which elements of the quality assessment checklist are essential for an article to pass the quality assessment step and move forward to the data extraction step. One method to determine which articles are of sufficient quality is to assign numerical values to checklist items to create a scale. A cut-point value for inclusion / exclusion is determined *a priori*. However, scales may mot allow the direction of any bias to be determined. Therefore, the use of individual components of methodological quality rather than summary scores is preferable (CRD, 2001). In our example checklists, some individual elements appear in '**bold**' in the checklists and those are considered essential to give a study a final 'yes' rating. An advantage of using this technique within the ESR software is that the criteria can be modified to increase or decrease the rigour of the quality assessment based on the quantity of literature that is available. If, for instance, only a small number of articles pass the screening step, the rigour can be relaxed by removing the "bold" designation from some of the less essential criteria. The software will then re-calculate the list of articles that pass the quality assessment step.

4.2 Use of quality assessment tools

Full articles are obtained for each of the references that pass the relevance screening step. An initial question in the quality assessment tool ascertains the type of study design. Based on the response to this question, the reviewer is directed to the appropriate checklist. Reviewers should be familiar with the process of critical review of the literature (see Appendix 8 for useful references on critical review).

The same considerations as those made during the process of selecting abstracts should be made when assessing the quality of the studies. At least two reviewers should assess the quality of each study. The ESR software allows reviewers to be assigned specific studies to review and can be used to store pdf files of articles. Each reviewer reads and scores each article independently using the appropriate quality assessment checklist. If none of the reviewers select exclusion responses, the reference will pass to the next level. If the reviewers agree on at least one of the exclusion responses, then the reference will be excluded. However, if none of the above criteria are met then the reference will stay in a state of conflict. Conflicts should be resolved by consensus between the two reviewers and there should be a procedure for identifying and resolving disagreements.

It is important to evaluate the quality assessment forms by using a pilot sample of articles to ensure that multiple reviewers can consistently apply the appraisal criteria. A useful validation and reviewer training strategy is to have all of the reviewers conduct quality assessment on the first 10 articles that pass relevance screening. Poor agreement between assessors may lead to revisions or rewording of quality assessment questions, or may lead to additional training of reviewers in critical review of the literature.

Many articles may be excluded at this stage. It is essential to include in the final report a table or flow chart with the list of excluded studies and the reason for excluding each article. A clear description of the quality assessment tool should also be documented.

5. Data Extraction

KEY POINTS

- **The objective of the data extraction form is to extract relevant information and results from studies that passed quality assessment in order to synthesize the results of studies included in the systematic review.**
- **The information extracted includes descriptive data to provide context, study characteristics that provide information to examine and explain heterogeneity between studies, and study results.**

The objective of the data extraction stage is to accurately extract information on relevant features and results of the selected studies. The data extracted is linked to the review question and planned assessment of the included studies and is the data repository from which data summarization or analysis is conducted (Cochrane, 2004). Because each review is different, data collection forms will vary across reviews. However, there are similarities in the types of information that are important, and forms can be adapted from one review to the next. It is helpful when preparing the data extraction form to have formulated the structure of the tables and figures that will be used to present the results so that the necessary data can be visually presented.

5.1 Developing the data extraction form

Reviewers should consider how much information they want to collect *a priori*. When deciding on the content of the form, reviewers should consider the information that will be needed to construct tables that summarize the studies included in the review and the data required from each study to perform analyses (Section 6). Forms that are too detailed can be wasteful of time. Conversely, if forms omit essential data, the reviewers may have to re-do data extraction steps which can be frustrating and time consuming. The data extraction form collects data for several purposes; descriptive data to provide context, describe heterogeneity in the data and allow stratification of summarized data, and research findings (results).

The section 'Data Extraction Forms' includes an example data extraction form designed for the review on "The use of probiotics to reduce *E. coli* O157 in the feces of post-weaned ruminants". The data extraction form contains the following sections:

1. General information
2. Population characteristics
3. Intervention protocols
4. Outcome measurement
5. Statistical analysis
6. Results

1. General Information

Data extraction forms can be adapted to different reviews and, in some cases, reviewers participate in multiple reviews (Cochrane, 2004). For this reason, the form should include a reference identification number for each study, bibliographic details of the paper from which the data are being extracted and the name of the reviewer who is abstracting data. It is important to record the source of information, especially when data were obtained from multiple reports of the same study. In addition, information on the language of publication and funding sources for the study should be recorded.

Information on the type of study design is included in the form. Different study designs provide different levels of evidence and therefore it is advisable to stratify the results by study type.

2. Population characteristics

The information collected in this section is used to describe the populations used in each study to provide context and to allow the reader to determine the external validity of the results for the purpose for which they will use them. Additionally, information is collected on population factors that may influence treatment effects and thereby may be necessary to explain and evaluate heterogeneity of study results (Cochrane, CRD, 2001). Therefore, data collected on population characteristics may include questions related to the country in which the study was conducted, selection of herds and the type of herds used, and attributes of the animals such as species, type of production, breed, age, and gender. In observational studies, some of these factors also may be considered as potentially confounding variables.

3. Interventions protocols

Details of the intervention and how it was administered to the population should be included because differences between intervention protocols may be an important source of heterogeneity among studies. Information on type of treatment, routes of delivery (such as orally through diet vs rumen canula), doses, frequency and length of administration should be recorded. In challenge trials, it is important to consider the level of challenge of the pathogen of interest, whether the animals tested negative to the pathogen of interest prior to inoculation, and whether animals were challenged before or during the intervention administration. It is also important to record what was given to the control group(s). Different types of control groups can be a source of differences among studies. Control groups may include non-treated controls, placebo-treated controls, or controls given an alternative treatment protocol.

Another feature that is relevant in farm-level research is the level at which treatments were allocated. In agriculture, grouping of animals is common in commercial settings and treatments may be allocated at the farm, pen, or individual animal level. Recording this information is important because it may pertain to external validity and practicality

of the intervention protocol, but also because it is essential to the statistical analysis and interpretation of the results.

4. Outcome measurements

Reports of studies often include more than one outcome. The specific question used for the systematic review will define the primary outcome. In the example of the use of probiotics to reduce *E. coli* O157 in ruminants, the primary outcomes were the prevalence or level of shedding of *E. coli* O157 in feces. Extracting data on secondary outcome, when available, may provide insight into potentially harmful or beneficial effects of the intervention. In on-farm studies, secondary outcomes may include mortality, morbidity, average daily gain, average feed conversion, and feed gain ratio. It is important to note, however, that the search strategy may not have been designed to identify all studies that potentially address the secondary outcomes. For instance, in the probiotic example, studies which addressed only the relationship between probiotic treatments and animal performance would not have been selected due to the use of *E. coli* terms in the outcome component linked to the other components with the use of "AND".

Information on the details of the laboratory technique(s) used to measure the outcome of interest (such as culture, serological, biochemical or molecular tests), the use of positive and negative controls, and the diagnostic criteria are important for descriptive purposes and as a potential source of heterogeneity between studies. For instance, over the past decade, there have been significant changes in the diagnostic testing protocols used to detect *E. coli* O157 that have resulted in higher sensitivity to detect the pathogen. This will impact the prevalence of the organism detected in RCTs and observational studies, which will impact study power and confidence.

5. Statistical Analysis

There are important statistical aspects in agriculture and agri-food research studies that should be recorded in the data extraction form:

1. *The number of farms, pens, and animals included in the study.*
2. *The level at which statistical analyses were performed and whether clustering was accounted for, when applicable.* Livestock populations are grouped in a very different way compared to human populations. Animals often are housed in pens within barns within farms. This means that a hierarchy and non-independence of study subjects within groups exist and should be accounted for in the analysis. Particularly in the older literature, it is not uncommon for animal clustering to not be adequately accounted for in analysis. Therefore, the systematic review team must decide whether or not to include studies that do not account for these factors. This decision may relate to the number of high quality studies available to address the review question. If such studies are included, the potential bias in the reported p-values should be identified when presenting the results.
3. *Repeated measures.* In studies of long duration, results may be presented for several periods of follow-up. In our probiotics examples, it was common for the

shedding of *E. coli* O157:H7 to be reported at more than one point in time. Some studies report outcome measures for each treatment group which incorporate all time points, such as total number of events, overall mean, or a trend over time. In other cases, studies report on single time points without accounting for the increased probability of type I errors with repeated independent assessments of the experimental units over time. As with clustered data, the review team needs to decide whether to include studies where measurements were taken over time and the analysis did not adequately account for this aspect.

NOTE: In some instances, while clustering and repeated measures may not have been adequately addressed in the analysis, the raw data for each experimental unit may have been provided. There should be a mechanism in the data extraction form for recording this information so that post-hoc statistical analyses can be performed.

4. *Confounding.* Confounding is present when study groups being compared differ in the frequency of the outcome for reasons other than the exposure of interest, distorting the association of interest. This is of particular concern in observational studies where random allocation to treatment groups is not used. Factors that may confound results in veterinary and agri-food studies may include age, gender, breed and weight of the animals, geographical location of farms, or the time of the year when the samples where taken. It is important to record which potentially confounding variables were accounted for and the means by which they were controlled (exclusion, matching, or analytical control).

6. Results

Primary and secondary outcomes can be measured on a continuous scale or as categorical data (often dichotomous). If either type of data is potentially relevant to answer the systematic review question, the data extraction form can be configured to capture data of both types. This will require the use of separate data extraction tables for the two types of outcomes. When there are both primary and secondary outcomes, data extraction tables are provided for each type of data for each potential outcome measure. The results may correspond to univariate associations at a single time, multi-variable associations adjusted for confounding or prognostic variables, associations adjusted for hierarchical structure (clustering) of data, associations adjusted for repeated measures, or combinations of these. In each case, the information for each outcome should be collected in the format that it was reported and, if necessary, can be transformed in a subsequent step. The effect measures for continuous and dichotomous outcomes that can be extracted are summarized as follows:

1. Continuous data information:

- Number of experimental units for each treatment level
- Overall, least square or contrast means for each treatment level
- Mean differences from control
- Unit of results
- Lower/Upper 95% CI
- Standard error
- P value

2. Dichotomous data information:

- Number of positive experimental units per treatment group
- Proportion of positive experimental units per treatment group
- Total number of experimental units per treatment group
- Unit of results
- Odd Ratio (OR)
- Relative Risk (RR)
- Lower/ Upper 95% CI
- P value

NOTE: For repeated measures trials, decisions will need to be made about which outcomes to assess:

- Overall means of treatment groups or interactions over time
- An overall significance of the treatment effect (or interaction term) with details reported for time points with statistically significant differences or for a pre-selected time point for all trials (in which case, some trials may not include evaluation at the selected time point).
    - The results from the longest follow-up period from each trial. This may induce a lack of consistency across studies that will increase heterogeneity.

7. General Comments

There should be a space allocated for writing general notes about the study, any problems and / or additional information that might be important for summarizing the data.

There are two possible ways to create a data collection form: paper and electronic forms. The following table summarizes the pros and cons for paper and electronic forms (adapted from Cochrane, 2004, CRD, 2001).

Paper Forms	Electronic Forms
Pros	
Easier to design.	Data entry and data abstraction are done in one step. Faster and more efficient
Easier to change original forms.	Data can be directly downloaded into statistical programs if meta-analysis is being performed.
Comparison of extraction results from multiple reviewers of the same study is simple. One form can be use to mark and correct errors and disagreements.	
Cons	
Data may need to be converted into files compatible with data analysis.	Need to design, pilot and refine an electronic copy of the form before data entry. This may require computer programming knowledge (although software such as ESR provides user-friendly interface).
No electronic documentation for long-term storage of results.	Difficult to change original forms.
	More difficult to identify and address errors and disagreements, especially with open-ended responses.
	May not be compatible with programs used to generate and store the final review.

When designing a data extraction form, the structure should follow a logical sequence. The decision rules and coding of responses need to be determined *a priori* and should be as consistent and straightforward as possible. Instructions can be added directly on the data extraction form near the data field.

The use of the ESR software allows the development of data extraction forms for data entry and subsequently the extrapolation of data files into spreadsheet software files (Excel, Quatro Pro). These systems are compatible with computer statistical analysis programs (SAS, Stata). Information from the data extraction forms can be cross-referenced to relevance or quality assessment questions by matching on the reference identifiers. This is particularly useful when the systematic review question contains multiple component questions.

5.2. Use of the data extraction form

The data extraction form should be tested by several reviewers on a sample of studies to ensure that the data entry follows a logical order and that the instructions are clear. The

testing of the forms will identify data that are not needed or missing and coding instructions that are confusing. Thus, testing the forms will ensure that all of the required information is extracted in a uniform way (Cochrane, 2004; CRD, 2001).

Data extraction should be performed independently by at least two reviewers and the data extracted by these reviewers should be compared to improve reliability. Any disagreements should be discussed and resolved either by consensus among reviewers or by the participation of an additional reviewer. If financial and time factors do not allow duplicate data extraction, a single reviewer can perform data extraction, with a second reviewer checking the first reviewer's work (CRD, 2001). It is advisable that a record is kept of any disagreements, and the changes made to address them, thus providing a historical record of the decisions and refinements that occur during the review (CRD, 2001)

In some instances, there are multiple reports based on the same data. It is important to identify multiple reports of the same data/ study, where papers report accumulating numbers of participants or increasing length of follow up because it would be misleading to include results of several reports of the same study. Studies reporting a greater treatment effect are more likely to be duplicated, leading to bias in the estimates of effectiveness.

In cases where reports do not provide all of the information that needs to be extracted, such as in proceeding abstracts or current research reports, it may be necessary to contact the authors of the study to request for additional data (CRD, 2001). This can be done by requesting the specific information needed to complete the form, or by providing the primary researcher(s) with the data extraction form and requesting that they complete the form. In this case, some mechanism for quality control of data entry should be determined.

6. Data Synthesis

KEY POINTS

- **The objective of the data synthesis is to summarize results from primary studies of sufficient quality using qualitative and, when possible, quantitative methods.**
- **Data are grouped and tabulated in a way that will permit the reader to visualize similarities and the differences of the characteristics between the studies included in the review, and the level of evidence provided by each of the studies.**
- **Results can be presented in tables or graphs that highlight whether the effect of a treatment is consistent and effective.**

The objective of data synthesis is to summarise and combine the results of primary studies included in the review, through a descriptive synthesis and if possible through a quantitative method using statistical techniques such as meta-analysis. Q*ualitative* summarization includes the tabulation of study characteristics (population, intervention and outcome) and results. *Quantitative* synthesis (meta-analysis) includes the use of statistical methods for assessing heterogeneity in results and for generating pooled results**.** Meta-analysis can only be used when the study designs and outcome definitions among studies are sufficiently homogenous to be combined into one pooled estimate. This chapter focuses on qualitative synthesis of the studies. Meta-analysis is beyond the scope of this manual and can be reviewed in Greenland, 1998; Cochrane Reviewers Handbook 4.2.2., 2004 and Dohoo et al, 2003.

6.1. Descriptive summarization of results

6.1.1 Study characteristics.

It is useful to produce tabular and graphical summaries of the results of the primary studies included in the review. For the descriptive component, the information already collected during the data extraction phase is grouped and tabulated in a way that will permit the reader to visualize the similarities and differences in the demographic characteristics among the included studies and the level of evidence provided by each of the studies. This provides context and allows an assessment of external validity.

The first table in the results section should include the principal author and year for each study. The results should be grouped by study design, because this will be an important source of heterogeneity among studies and provides information on the level of evidence supplied by the results (see section 1.3). This table also includes information that succinctly describes population characteristics such as type and attributes of animals, their age, breed, and gender, the country where the subjects were located, and the setting where the study was performed. An example of the descriptive table layout (with hypothetical results to illustrate) is as follows:

Study	Country	Setting	Attributes of animals	Type of Animal	Age	Breed	Gender
RCTs							
Author 1 et al., 1997	USA	Single farm (University)	Commercial	Beef (feedlot)	Mature	Mixed	Mixed
Author 2 et al., 2004	USA	100 commercial farms	Commercial	Beef (feedlot)	Mature	British X	Steer
Challenge trials							
Author 3 et al., 2001	Canada	Single farm (University)	Gnotobiotic	Dairy	Pre-weaned	Holstein	Male
Cross-sectional observational							
Author 4 et al., 2003	Finland	40 commercial farms	Commercial	Dairy	Mature	Holstein	Female

6.1.2 Intervention.

A table should be provided with the intervention(s) and intervention protocols with dose, route of administration, unit of allocation, number of animals and pens used per treatment group, and type of controls. A final column in this table may be used to provide an overall conclusion (e.g., protective, no difference, positive effect). In many pre-harvest agri-food intervention studies, especially in randomized controlled trials and challenge trials, more than one treatment group may be included. In the case of the example probiotic study, some of the trials included multiple treatment groups and / or multiple probiotic products within a treatment group. The description of the interventions should be grouped within study and within study design. The assessment of the information included in the tables is important to highlight the comparisons that were made. An example of the intervention table layout (with hypothetical results to illustrate) is as follows:

Study	Intervention	# Farms	# Pens	# Animals	Unit of allocation	# Units / tx group	Route	Dose	Results
RCTs									
Author 1 et al., 1997	Tx A vs. placebo	1	20	200	Pen	10	Oral	100 g/kg	NS
Author 2 et al., 2004	Tx B vs. placebo	100	400	4000	Pen	200	Injection	10 g/kg	Pos.
	Tx. B&C versus placebo	100	400	4000	Pen	200	Injection	5 g/kg	NS

6.2 Summarization of study results

There are several types of data commonly reported in primary studies: Dichotomous or binary data where each individual is classified in only one of two possible values and continuous data which may take any value within a defined range. When dichotomous data are used, the results are expressed as odds ratios, risk ratios or risk differences. With continuous data, the results are summarized as means, mean differences, or standardized means (Table 7.1). Studies also may use survival or time to event data where the outcome is the time to the occurrence of an event and the results are reported using hazard ratios.

Table 6.1. Outcome measures in primary studies using continuous or binary data.

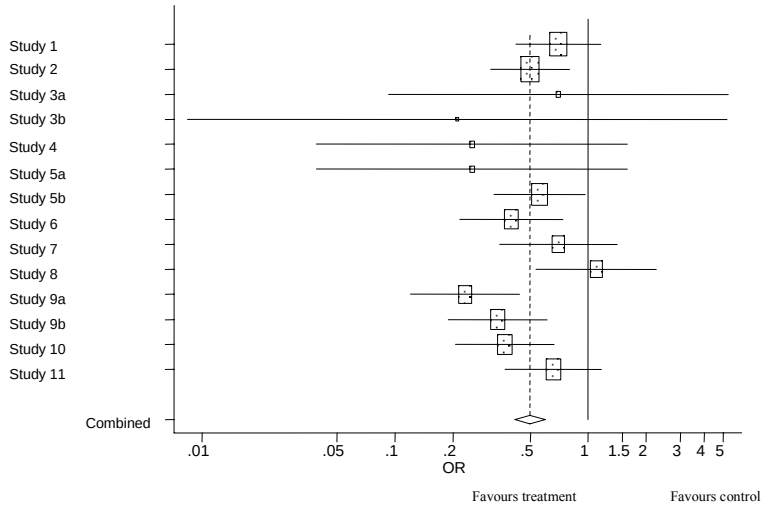
Outcome Measures	Description
Continuous outcomes	
Difference between group means (MD)	If the outcome of interest is continuous and the mean of the treatment and control are reported, then the difference between means can be calculated.
Standardized difference (SD)	Differences between means in each study are standardized by an estimate of the standard deviation. This removes the effect of the scale of measurement, but can be difficult to interpret.
Weighted difference in means (WD)	An average (pooled) difference between mean values for treatment and control groups can be calculated across a group of studies if they have used the same outcome measure.
Standardized weight difference (SWD)	An average (pooled) difference between treatment and control group mean values across a group of studies where the outcomes among studies were measured using different scales that cannot be converted to a common measure
Binary outcomes	
Risk difference (RD)	The absolute difference in risk between the treated and control groups indicates any differences in the probability of disease in a treated group, beyond the baseline risk, that results from treatment. An RD below 1 means that the intervention was protective.
Relative risk or Risk ratio (RR)	The risk of the outcome in the treatment group relative to that in the control group. An RR below 1 means that the intervention is protective
Odds ratio (OR)	The odds of a positive outcome in the treatment group relative to the corresponding odds in the control group. An OR below 1 means that the intervention is protective.

Adapted from Glasziou et al, 2001, CRD, 2001 and Cochrane, 2004

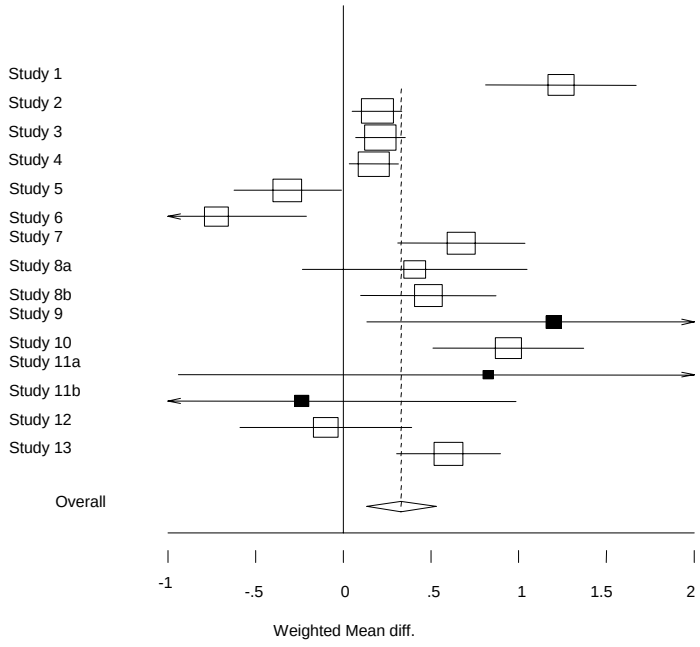
It is important to decide which effect measure(s) will be used to describe and summarise the data in the systematic review. In systematic reviews, data from multiple studies can only be combined quantitatively when the same outcome measure is used. In some instances, it may be possible to transform data to standardize outcome measures across studies.

Results from multiple studies may be summarized using a graphical approach called Forest plots. Forest plots use point estimates along with their confidence intervals and may help to reveal discernable patterns in the data among studies. A visual examination of the Forest plot gives an idea of the heterogeneity of the results between studies. Separate Forest plots may be used for studies with different study designs or different outcomes. Examples of Forest plots for different types of studies and outcomes, with hypothetical parameters and confidence intervals, are illustrated bellow.

Randomized Controlled Trials, dichotomous outcome.



Challenge trials, continuous data.



These graphs show hypothetical results categorized according to study design and outcome type. Results can be presented in different ways, such as by grouping studies by species, or by other important demographic or study design characteristics. The RCT Forest plot suggests that there might not be heterogeneity among treatment groups within this study type. The confidence intervals of all treatment groups overlap with the confidence intervals of the overall OR estimate (0.5). The box sizes illustrate the contribution of each treatment to the overall effect. The Forest plot for challenge trials indicates that there might be some heterogeneity among treatment groups. The investigation of the variation of the results within and among studies and study types and the sources of the variation are very important and should be investigated.

6.3. Heterogeneity

The investigation of differences between studies included in a review is an essential part of the summary of the data. There are several factors that can cause the variation in the effects observed among studies including study design, differences in intervention protocols, attributes of populations or individual animals, and differences in the outcome measurement. Forest plots provide a qualitative method of visualizing these differences.

References

AHRQ (Agency for Healthcare Research Quality). West, S., King, V., Carey, T.S., Lohr, K.N., McKoy, N., Sutton, S.F. and Lux, L. (Research Triangle Institute-University of North Carolina Evidence-based Practice Center Research Triangle Park, North Carolina): Systems to Rate the Strength of Scientific Evidence. (2002) Evidence Report / Technology Assessment No. 47.

CRD (Centre for Reviews and Dissemination) Report Number 4: Undertaking Systematic Reviews on Effectiveness. CRD'S Guidance for those carrying out or commissioning reviews. (2nd edition) (2001) NHS Centre for Reviews and Dissemination, University of York, York.

Chalmers TC, Smith H Jr, Blackburn, et al. A method for assessing the quality of randomized control trials. (1981) Control Clin Trials, 2:31-49.

Cochrane Collaboration: Cochrane Reviewers' Handbook 4.2.2. (Updated March 2004).

Cook, D.J., Mulrow, C.D., Haynes, R.B. Systematic reviews: Synthesis of best evidence for clinical decisions. Annals of Internal Medicine (1997), 126 (5) 376-380.

Dickersin K. How important is publication bias? A synthesis of available data. AIDS Education and Prevention (1997), 9 Suppl A:15-21.

Dohoo, I., Martin, S.W., Stryhn, H. Meta-analysis. *In:* Veterinary Epidemiology (2003) AVC Inc., University of Prince Edward Island, Canada. pp. 543-560.

Ferguson, J., Dalton, C., McGettigan, P. and Hill, S.: Antibiotic Resistance in Animal Enteric Bacteria and Human Disease – a Review of the Scientific Literature. Prepared for the Joint Expert Technical Advisory Committee on Antibiotic Resistance (JETACAR) (December 1998). The University of Newcastle, Australia.

Glasziou, P., Irwig, L., Bain, C. and Colditz, G.: Systematic reviews in health care: A practical guide (2001) University of Cambridge, United Kingdom.

Greenland, S. Meta-analysis. *In*: Modern Epidemiology. (1998) Rothman, J.K. and Greenland S *eds*, 2nd edition. Lippincott Williams & Wilkins pp 643-673.

Haynes RB, Wilczynski N, McKibbon, KA, Walker, CJ and Sinclair JC: Developing optimal search strategies for detecting clinically sound studies in Medline. (1994) Journal of the American Medical Information Association, 1: 447-458.

Mead, P.S., Slutsker, L., Dietz, V., McCaig, L.F., Bresse, J.S., Shapiro, C. Griffin, P.M., Tauxe, R.V. Food-related illness and death in the United States. (1999) Emerging Infectious Diseases. 5, 607-625.

NHMRC (National Health and Medical Research Council): How to review the evidence: systematic identification and review of the scientific literature. Handbook series on preparing clinical practice guidelines (1999) Legislative Services, AusInfo, GPO Box 1920, Canberra ACT 2601, Australia.

O'Blenis P and Garritty C. The Electronic Systematic Review Handbook : Practical Concepts and Methods for Electronic Screening and Data Abstraction. (2004) First Trial Stat Corporation Paperback Edition, Canada.

Appendix 1

Key distinctions between narrative and systematic reviews, by core features of such reviews

Core feature	Narrative Review	Systematic Review
Study question	Often broad in scope.	Often a focused question.
Data sources and search strategy	Details of the databases searched and the search strategy are not typically provided.	Comprehensive search of electronic databases and unpublished data. Explicit search strategy provided.
Selection of articles for study	Not usually specified, potentially biased.	Criterion-based selection uniformly applied.
Article review or appraisal	Variable, depending on who is conducting the review.	Rigorous critical appraisal, typically using a quality assessment form.
Study quality	If assessed, may not use formal quality assessment.	Some assessment of quality is almost always included as part of the data extraction process.
Synthesis	Often a qualitative summary.	Quantitative summary (meta-analysis) if the data can be appropriately pooled; qualitative otherwise.
Inferences	Sometimes evidence-based.	Usually evidence-based.

Sources: Adapted from Cook et al., 1997

Appendix 2

Table 1. Population, intervention and outcome terms used for a systematic review on the use of probiotics for the reduction of *E. coli* O157:H7 in the feces of beef and dairy cattle, with the number of articles retrieved in PubMed and Agricola databases.

	PubMed	Agricola		PubMed	Agricola
Population			**Intervention**		
Bovine	265,289	10,000*	Probiotic	1,888	1,059
Cattle	228,782	10,000*	Probiotic culture	292	54
Beef	10,851	10,000*	Probiotic bacteria	1,602	117
Cow	221,838	10,000*	Selected lactic acid bacteria	227	12
Cows	220,890	10,000*	Lactic acid bacteria	3,968	3,212
Steer	2,099	7,464	Lactic acid-producing bacteria	358	30
Steers	2,827	6,339	Lactic acid bacteria mixture	128	1
Heifer	1,182	7,345	Lactic fermentation	1,202	133
Calf	29,581	7,990	Lactic fermentation product	341	1
Calves	15,474	10,000*	Direct fed microbials	10	21
Ruminants	308,712	5,998	Live bacteria supplements	27	0
Ruminant	308,755	10,000*	Competitive exclusion	651	229
Rum*	29,894	10,000*	*Lactobacillus*	12,094	4,791
Rumin*	12,455	1,277	*Bifidobacteria*	889	348
Beef steers	625	522	*Propionibacterium*	4,123	297
Beef calf	575	22	Lactobacillus fermentation products	341	1
Beef calves	952	436	Based direct-fed microbials	2	1
Beef herd	456	428	Strain mixture	2,078	104
Beef cattle	8,070	10,000*	Dietary supplementation	11,112	553
Beef farm	198	58	Yeast	128,327	10,000*
Beef cow	7,844	2,789	Lactic acid	22,433	6,439
Beef cows	7,841	2,525	Bacteriocin	3,148	1,131
Feedlot cattle	948	565			
Feedlot steers	367	243			
Dairy cattle	11,948	9,075			
Dairy herd	2,113	3,888			
Dairy cow	11,652	10,000*			
Dairy cows	11,941	10,000*			
Dairy calf	82	82			
Dairy claves	1,281	527			
Milk cows	18,498	209			
Calf herd	428	70			
Calf herds	426	33			
Calf cattle	13,310	1			
Dairy	19,285	10,000*			

*Search was truncated

Agricola includes: journal articles, book chapters, short reports and reprints

	PubMed	Agricola		PubMed	Agricola
Outcome			**Outcome**		
Escherichia coli	217,670	10,000*	Bacteria load	7,260	0
Escherichia coli O157	3,008	1,197	Loads of bacteria	7,260	1
Escherichia coli O157:H7	2,050	1,044	Level of bacteria	45,099	22
E. coli	181,725	4,664	Pathogen level	2,323	11
E. coli O157	2,828	806	Pathogen concentration	1,870	7
E. coli O157:H7	1,890	701	Bacteria concentration	65,045	13
Enterohaemorragic	246	0	Bacteria level	45,099	15
Enterohaemorragic *E.coli*	2	0	CFU	17,591	3,072
Enterohemorragic *Escherichia coli*	7	2	CFU per gram	339	24
Verotoxigenic	96	40	CFU per ml	2,585	18
Verotoxigenic *E. coli*	89	9	Colony forming units	118,672	686
Verotoxigenic *Escherichia coli*	95	31	Colony forming units per gram	262	32
Shiga-like toxin	2,262	110	Colony forming units per ml	5,186	7
Shiga-like toxin *E. coli*	1,753	0	Bacteria logs	487	0
Shiga-like toxin *Escherichia coli*	1,893	0	Bacteria fecal count	568	0
EHEC	604	84	Bacteria log count	1,283	0
VTEC	432	66	Bacteria log counts	739	0
STEC	745	886			
Enteric	17,984	1,829	Fecal	22,832	3,894
Coliform	2,130	1,795	Feces	59,322	5,996
Coliforms	1,844	841	Manure	1,880	10,000*
Faecal shedding	498	7	Rumen fluid	1,289	1,350
Fecal shedding	492	98	Gastrointestinal	163,423	146
Fecal prevalence	3,054	1	Pre-harvest	107	397
Fecal recovery	740	15	Gastro-intestinal	4,950	521
Fecal incidence	3,280	0	Preharvest	153	947
Fecal level	1,631	3			
Fecal persistence	136	1			
Fecal isolation	4,068	1			
Fecal detection	1,673	1			
Fecal inhibition	359	0			
Fecal recovery	740	15			
Fecal reduction	1058	0			
Fecal carriage	120	3			
Fecal elimination	636	11			

*Search was truncated

Agricola includes: journal articles, book chapters, short reports and reprints

Table 2. Number of abstracts obtained in PubMed using different combinations of population terms

Population terms	Number of papers PubMed
Bovine	271,270
Cattle	232,165
Beef	11,088
Dairy	19,752
Ruminant	312,695
Bovine or Cattle	277,357
Bovine of Beef	274,205
Bovine or Dairy	278,770
Bovine or Ruminant	359,928
Bovine or Cattle or Beef	280,126
Bovine or Cattle or Dairy	284,620
Bovine or Cattle or Ruminant	365,324
Bovine or Cattle or Beef or Ruminant	368,047
Bovine or Cattle or Beef or Dairy or Ruminant	374,517

Appendix 3

Table 1. General and specific terminology used in different agricultural populations and in food safety

	Swine	Bovine	Poultry	Food Safety
General terminology	Pig, pork, porcine, herd, farm	Ruminants, cows, heifer, herd, cattle, beef, farm, buiatric	Avian, chicken, bird, farm	Food chain, food risk, food products, food-born disease, food-borne illness, food-borne infection.
Specific terminology	Suckling, nursing, pre-weaned, weanling, weaned, nursery, post-weaning, early weaning, growing-finisher, growing-finishing, finisher, slaughter, lactating sow, dry sow, gilts, hogs, barrow, boar, reproductive boars, farrow-to finish, farrow-to-wean, segregated early weaning, off-site nursery, pot-bellied, porcine industry, swine industry.	Dairy cows, milk cow, lactating cow, dry cow, dairy calves, dairy herd, dairy farm, dairy cattle, cattle herd, beef herd, beef cattle, beef farm, beef cows, calf, calves, calf herd, calf cattle, neonatal calves, pre-weaned calves, post-weaned calves.	Chicks, broilers, domestic birds, growing chicken, growing chicks, chicken flocks, growing broilers, hens, layer hens, growing laying hens, fowls, domestic fowls, pullet.	Diet, meat, beef, pork, poultry, veal, milk, dairy, egg, cheese products, ground beef, ground meet, veal, sea-food, seafood toxins, food-borne parasites Food-borne, food-borne, food animals, food producing animals, food safety risk, water-borne, boiled or cooked vegetables, fresh vegetables, food related allergies, travel related foodborne diseases, communicable diseases, restaurant, grocery store, slaughterhouse, fast food, antibiotic resistance.

Table 2. General and specific terminology used for different interventions in agri-food public health

	Prevention	Therapeutic	Risk factors
General terminology	Control, surveillance, strategy, programs, prophylaxis, screening, prevention, outbreak investigation, reservoirs, transmission patterns.	Treatment, drug, intervention, therapy.	Putative causal factor, agent factors, non-agent factors, risk assessment, risks.
Specific terminology	Growth promoters, vaccination, biosecurity, alternative medicine, diet supplements, diet, sanitizers, food safety risk assessment, screening of bacteria, quality control.	Chemical agent, antimicrobials, irradiation, surgery, alternative medicine, chemotherapy, dietary supplements, dietary change.	Host, population, immunity, genetic, environmental factors, environmental pollution, on farm factors, temperature contamination, biosecurity, hygiene, management and diet changes, texture of feed, transportation, food packaging, bacteria loads, level of toxins, parasites loads, hazardous substances, chemical contamination, food handle, food preparation, preparation methods, climate change, transmission pattern.

Table 3. General and specific terminology used for different outcomes in agriculture and food safety

	Agriculture	Food Safety
General terminology	Production, profit, animal health, productivity, profitability, improvement, performance, development, economic impact, wealth, significance, prevalence, risk, incidence, rates, mortality, morbidity, efficiency, bacteria load, pathogen level, body condition, quality.	Food and water outbreak, infection, intoxication, disease, illness, public health, significance, prevalence, risk, incidence, rates, bacteria load, pathogen level, quality, impact, economic impact.
Specific terminology	Weight gain, average daily gain, average daily feed intake, feed efficiency, feed gain ratio, growth performance, breed efficiency, reproductive efficiency, cost benefit, price-cost, treatment cost, culling rate, death rate, kilograms produced of meat, length of productive time. **Swine:** Pigs weaned per sow per year, litter per sow per year, pigs weaned per litter, average weaning age, average weaning weight, pigs born alive per litter, pigs weaned per crate per year, pigs weaned per lifetime female, non-productive days, average lactation length, farrowing rate, weaning-to-breed interval, days to market, feed cost per animal, feed cost per unit of gain, sow boar ratio. **Bovine:** Milk production, days in milk, daily milk yield, calving interval, calf sold per year, number of calves weaned. **Poultry:** Kilograms of broiler per square meter, egg production rates, egg shell quality.	

Appendix 4

Number of abstracts obtained for various population search terms by livestock species from PubMed, Agricola and CABI

	PubMed	Agricola	CABI		PubMed	Agricola	CABI
Swine	125,184	10,000+	42,678	**Bovine**	265,289	10,000+	81,696
Pigs	243,695	10,000+	134,464	Cattle	228,782	10,000+	338,889
Pork	2,428	6,932	5,856	Beef	10,851	10,000+	31,725
Porcine	130,069	7,201	21,206	Cow	221,838	10,000+	41,843
Suckling pig	1,199	307	28	Cows	220,890	10,000+	168,645
Nursing pig	280	66	13	Steer	2,099	7,464	2,350
Nursing swine	246	1	0	Steers	2,827	6,339	10,460
Weanling pig	571	283	92	Heifer	1,182	7,345	3,846
Weaner pig	99	49	48	Calf	29,581	7,990	30,741
Weaned pig	1256	827	133	Calves	15,474	10,000+	52,547
Nursery pig	215	92	47	Ruminants	308,712	5,998	442,869
Post-weaning pigs	454	10	49	Ruminant	308,755	10,000+	8,902
Early-weaning pigs	327	13	30	Rum*	29,894	10,000+	1,411
Growing-finishing pigs	285	312	1,083	Rumin*	12,455	10,000+	162
Grower-finisher pigs	79	19	87	Beef steers	625	522	708
Growing hog	18	4	0	Beef calf	575	22	52
Finishing pigs	643	565	1,980	Beef calves	952	436	768
Finisher pigs	130	41	204	Beef herd	456	428	261
Finishing hogs	17	2	18	Beef cattle	8,070	10,000+	11,360
Sows	3,333	5,008	19,940	Beef farm	198	58	73
Gilts	2,377	2,512	8,309	Beef cow	7,844	2,789	407
Hogs	433	745	318	Beef cows	7,841	2,525	2,497
Boars	120,471	2,332	8,525	Feedlot cattle	948	565	665
Boar	2,303	9,117	5,863	Feedlot steers	367	243	287
Barrow	2,106	1,115	437	Dairy cattle	11,948	9,075	20,397
Slaughter pigs	1,493	190	1,132	Dairy herd	2,113	3,888	3,266
Pig farm	1,718	2,491	833	Dairy cow	11,652	10,000+	2,902
Pig herd	1,178	158	282	Dairy cows	11,941	10,000+	23,555
Swine farm	1,687	569	75	Dairy calf	82	82	99
Swine herd	1,182	289	148	Dairy claves	1,281	527	985
Porciculture	3	0	0	Milk cows	18,498	209	2,611
Porcine industry	2,998	0	0	Calf herd	428	70	70
Swine industry	3,137	166	91	Calf herds	426	33	65
Farrow-to finish	97	97	249	Calf cattle	13,310	1	26
Segregated early weaning	26	18	50				
Multi-site	681	52	9				

+Search was truncated
Agricola includes: journal articles, book chapters, short reports
CABI: CAB direct international

	PubMed	Agricola	CABI
Poultry	97,231	10,000+	123,042
Avian	138,714	10,000+	29,477
Chicken	92,102	10,000+	22,814
Bird	127,541	4,369	9,821
Chicks	10,738	9,188	20,342
Broilers	2,778	7,755	16,655
Domestic bird	4,915	8	20
Domestic birds	5,042	96	304
Growing chicken	1,446	30	44
Growing chicks	464	250	318
Chicken flocks	1,153	88	293
Growing broilers	195	77	154
Hen	7,556	3,170	8,626
Hens	5,737	8,227	18,282
Layer hens	390	48	158
Growing laying hens	67	2	0
Fowls	1,456	3,925	24,632
Fowl	4,955	550	89,063
Domestic fowl	92,425	1,336	2,031
Domestic fowls	92,190	114	529
Pullet	185	204	488

+Search was truncated
Agricola includes: journal articles, book chapters, short reports
CABI: CAB direct international

Appendix 5

Search terms strategy used in different databases for the project "Effect of using probiotics on the reduction of *E. coli* O157 in the faeces of post-weaned ruminants"

Overview of search terms by component:
Population: Ruminant(s), bovine, cattle, cow, steer, cow, calf, calves, beef, farm, herd, sheep, goat, deer, lamb(s)
Intervention: Probiotic*, lactobac*, bifidobac*, propionibac*, saccharomyces, fermentation, yeast, bacteriocin, competitive exclusion, strain mixture, dietary supplementation, lactic acid
Outcome: *Escherichia coli*, O157, O157:H7, bacteria load, bacterial load, bacteria counts, bacterial counts, bacteria log, bacterial log, bacteria level, bacterial level, feces, faeces, fecal, faecal, gastrointestinal, manure, coliform, enteric, ehec, shiga-like toxin

Databases and Search terms

A. Silverplatter:
Webspirs (**Medline, Current Contents, CAB Health, FSTA Retrospective, FSTA (1990-2005 U of Guelph), Pubmed**

(ruminant* or bovine or cattle or cow* or steer or calf or calves or beef or farm or herd or sheep or goat* or deer or lamb*[use lamb or lambs instead of lamb* in PubMed])

AND

(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin or "competitive exclusion" or "strain mixture" or "dietary supplementation" or "lactic acid")

AND

("escherichia coli" or o157 or "o 157" or "bacteria load" or "bacterial load" or "bacteria counts" or "bacterial counts" or "bacteria log" or "bacterial log" or "bacteria level" or "bacterial level" or feces or faeces or fecal or faecal or gastrointestinal or manure or coliform or enteric or ehec or "shiga-like toxin")

NOT

(milk or cheese or yogurt or yogourt or yoghurt or yoghourt)

B. CSA:
Biological Sciences (journals, conferences, websites, books, reports, dissertations, patents)

no quotation marks: words next to each other automatically treated as a phrase; use all words for plural or use ? for one character at the end or middle of word

ruminant? or bovine or cattle or cow? or steer or calf or calves or beef or farm or herd or sheep or goat? or deer or lamb?

AND

probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin or competitive exclusion or strain mixture or dietary supplementation or lactic acid

AND

escherichia coli or o157 or o 157 or bacteria load or bacterial load or bacteria counts or bacterial counts or bacteria log or bacterial log or bacteria level or bacterial level or feces or faeces or fecal or faecal or gastrointestinal or manure or coliform or enteric or ehec or shiga-like toxin

NOT

milk or cheese or yogurt or yogourt or yoghurt or yoghourt

C. CAB Direct (U of G)
no quotation marks: words next to each other automatically treated as phrase
use + at end of word to get variations on word or use ? for one character at the end or middle of word

ruminant? or bovine or cattle or cow? or steer or calf or calves or beef or farm or herd or sheep or goat? or lamb?

AND

probiotic? or lactobac? or bifidobac? or propionibac? or saccharomyces or fermentation or yeast or bacteriocin or competitive exclusion or strain mixture or dietary supplementation or lactic acid

AND

escherichia coli or o157 or o 157 or bacterial load or bacterial counts or bacterial log or bacterial level or feces or faeces or fecal or faecal or gastrointestinal or manure or coliform or enteric or ehec or shiga-like toxin

NOT

milk or cheese or yogurt or yogourt or yoghurt or yoghourt

D. EMBASE

bovids or cattle or beef cattle or dairy cattle or cow or sheep or goat or lamb or farm or herd

AND

probiotic or probiotic agent or lactobacillus of bifidobacterium or propionibacterium or saccharomyces or anaerobic fermentation or bacteriocin or lactic acid bacterium or strain mixture

AND

escherichia coli or escherichia coli O157 or verotoxin or feces or faeces or fecal or faecal or gastrointestinal or manure of coliform or enteric or bacteria count or bacteria load or bacteria log or bacteria level

E. UMI Proquest Dissertations (University of Guelph

(ruminant or bovine or cattle or cow or steer or calf or calves or beef or farm or herd or sheep or goat or deer or lamb

AND

(probiotic or lactobacillus or bifidobacterium or propionibacterium or saccharomyces or fermentation or yeast or bacteriocin or "competitive exclusion" or "strain mixture" or "dietary supplementation" or "lactic acid")

AND

("escherichia coli" or o157 or "o 157" or "bacterial load" or "bacterial counts" or "bacterial log" or "bacterial level" or feces or faeces or fecal or faecal or gastrointestinal or manure or coliform or enteric or ehec or "shiga-like toxin")

NOT

(milk or cheese or yogurt or yogourt or yoghurt or yoghourt)

F. Agricola

NB: Based on the structure of agricola for search term entry, the following would represent the necessary search term format. However, agricola does not allow this many search terms to be simultaneously entered. Therefore, for this example, it was necessary to enter the search terms in a series of smaller combinations (labelled 1 to 31, below).

((ab=(ruminant or ruminants or ruminant* or bovine or cattle or cow or cows or steer or calves or calf or beef or farm* or herd* or sheep or goat or goats or deer or lamb or lambs)) or (ti=(ruminant or ruminants or ruminant* or bovine or cattle or cow or cows or steer or calves or calf or beef or farm* or herd* or sheep or goat or goats or deer or lamb or lambs)) or (ke=(ruminant or ruminants or ruminant* or bovine or cattle or cow or cows or steer or calves or calf or beef or farm* or herd* or sheep or goat or goats or deer or lamb or lambs))) and ((ab=(probiotic* or probiotics or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin or "competitive exclusion" or "strain mixture" or "dietary supplementation" or "lactic acid")) or (ti=(probiotic or probiotics or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin or "competitive exclusion" or "strain mixture" or "dietary supplementation" or "lactic acid")) or (ke=(probiotic or probiotics or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin or "competitive exclusion" or "strain mixture" or "dietary supplementation" or "lactic acid"))) and ((ab=(feces or faeces or fecal or faecal or manure or gastrointestinal or coliform or enteric or ehec or o157 or "o 157" or "shiga-like toxin" or "escherichia coli" or "bacteria load" or "bacterial load" or "bacteria counts"or "bacterial counts" or "bacteria log" or "bacterial log" or "bacteria level" or "bacterial level") or (ti=(feces or faeces or fecal or faecal or manure or gastrointestinal or coliform or enteric or ehec or o157 or "o 157" or "shiga-like toxin" or "escherichia coli" or "bacteria load" or "bacterial load" or "bacteria counts"or "bacterial counts" or "bacteria log" or "bacterial log" or "bacteria level" or "bacterial level") or (ke=(feces or faeces or fecal or faecal or manure or gastrointestinal or coliform or enteric or ehec or o157 or "o 157" or "shiga-like toxin" or "escherichia coli" or "bacteria load" or "bacterial load" or "bacteria counts"or "bacterial counts" or "bacteria log" or "bacterial log" or "bacteria level" or "bacterial level"))) not ((ab=(milk or cheese or yogurt or yogourt or yoghurt or yoghourt)) or (ti=(milk or cheese or yogurt or yogourt or yoghurt or yoghourt)) or (ke=(milk or cheese or yogurt or yogourt or yoghurt or yoghourt)))

Agricola Search Combinations

1.
((ab=(ruminant* or farm or herd)) or (ti=(ruminant* or farm or herd)) or (ke=(ruminant* or farm or herd))) and ((ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) or (ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) or (ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin))) and ((ab=(feces or faeces or fecal or faecal or manure)) or (ti=(feces or faeces or fecal or faecal or manure)) or (ke=(feces or faeces or fecal or faecal or manure)))

2.
((ab=(ruminant* or farm or herd)) or (ti=(ruminant* or farm or herd)) or (ke=(ruminant* or farm or herd))) and ((ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) or (ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) or (ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin))) and ((ab=(gastrointestinal or enteric)) or (ti=(gastrointestinal or enteric)) or (ke=(gastrointestinal or enteric)))

3.
((ab=(ruminant* or farm or herd)) or (ti=(ruminant* or farm or herd)) or (ke=(ruminant* or farm or herd))) and ((ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) or (ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) or (ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin))) and ((ab=(coliform or ehec or o157)) or (ti=(coliform or ehec or o157)) or (ke=(coliform or ehec or o157)))

4.
((ab=(ruminant* or farm or herd)) or (ti=(ruminant* or farm or herd)) or (ke=(ruminant* or farm or herd))) and ((ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) or (ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) or (ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin))) and ((ab=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157)) or (ti=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157)) or (ke=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157)))

6.
(ab=(ruminant* or farm or herd) or ti=(ruminant* or farm or herd) or ke=(ruminant* or farm or herd)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(feces or faeces or fecal or faecal or manure) or ti=(feces or faeces or fecal or faecal or manure) or ke=(feces or faeces or fecal or faecal or manure))

7.
(ab=(ruminant* or farm or herd) or ti=(ruminant* or farm or herd) or ke=(ruminant* or farm or herd)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(gastrointestinal or enteric) or ti=(gastrointestinal or enteric) or ke=(gastrointestinal or enteric))

8.
(ab=(ruminant* or farm or herd) or ti=(ruminant* or farm or herd) or ke=(ruminant* or farm or herd)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(coliform or ehec or o157) or ti=(coliform or ehec or o157) or ke=(coliform or ehec or o157))

9.
(ab=(ruminant* or farm or herd) or ti=(ruminant* or farm or herd) or ke=(ruminant* or farm or herd)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157) or ti=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157) or ke=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157))

11.
(ab=(bovine or cattle or cow* or steer or calf or calves or beef) or ti=(bovine or cattle or cow* or steer or calf or calves or beef) or ke=(bovine or cattle or cow* or steer or calf or calves or beef)) and (ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) and (ab=(feces or faeces or fecal or faecal or manure) or ti=(feces or faeces or fecal or faecal or manure) or ke=(feces or faeces or fecal or faecal or manure))

12.
(ab=(bovine or cattle or cow* or steer or calf or calves or beef) or ti=(bovine or cattle or cow* or steer or calf or calves or beef) or ke=(bovine or cattle or cow* or steer or calf or calves or beef)) and (ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) and (ab=(gastrointestinal or enteric) or ti=(gastrointestinal or enteric) or ke=(gastrointestinal or enteric))

13.
(ab=(bovine or cattle or cow* or steer or calf or calves or beef) or ti=(bovine or cattle or cow* or steer or calf or calves or beef) or ke=(bovine or cattle or cow* or steer or calf or

calves or beef)) and (ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) and (ab=(coliform or ehec or o157) or ti=(coliform or ehec or o157) or ke=(coliform or ehec or o157))

14.
(ab=(bovine or cattle or cow* or steer or calf or calves or beef) or ti=(bovine or cattle or cow* or steer or calf or calves or beef) or ke=(bovine or cattle or cow* or steer or calf or calves or beef)) and (ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) and (ab=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157) or ti=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157) or ke=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157))

16.
(ab=(bovine or cattle or cow* or steer or calf or calves or beef) or ti=(bovine or cattle or cow* or steer or calf or calves or beef) or ke=(bovine or cattle or cow* or steer or calf or calves or beef)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(feces or faeces or fecal or faecal or manure) or ti=(feces or faeces or fecal or faecal or manure) or ke=(feces or faeces or fecal or faecal or manure))

17.
(ab=(bovine or cattle or cow* or steer or calf or calves or beef) or ti=(bovine or cattle or cow* or steer or calf or calves or beef) or ke=(bovine or cattle or cow* or steer or calf or calves or beef)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(gastrointestinal or enteric) or ti=(gastrointestinal or enteric) or ke=(gastrointestinal or enteric))

18.
(ab=(bovine or cattle or cow* or steer or calf or calves or beef) or ti=(bovine or cattle or cow* or steer or calf or calves or beef) or ke=(bovine or cattle or cow* or steer or calf or calves or beef)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1

acid)) and (ab=(coliform or ehec or o157) or ti=(coliform or ehec or o157) or ke=(coliform or ehec or o157))

19.
(ab=(bovine or cattle or cow* or steer or calf or calves or beef) or ti=(bovine or cattle or cow* or steer or calf or calves or beef) or ke=(bovine or cattle or cow* or steer or calf or calves or beef)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157) or ti=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157) or ke=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157))

21.
(ab=(sheep or goat* or deer or lamb*) or ti=(sheep or goat* or deer or lamb*) or ke=(sheep or goat* or deer or lamb*)) and (ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) and (ab=(feces or faeces or fecal or faecal or manure) or ti=(feces or faeces or fecal or faecal or manure) or ke=(feces or faeces or fecal or faecal or manure))

22.
(ab=(sheep or goat* or deer or lamb*) or ti=(sheep or goat* or deer or lamb*) or ke=(sheep or goat* or deer or lamb*)) and (ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) and (ab=(gastrointestinal or enteric) or ti=(gastrointestinal or enteric) or ke=(gastrointestinal or enteric))

23.
(ab=(sheep or goat* or deer or lamb*) or ti=(sheep or goat* or deer or lamb*) or ke=(sheep or goat* or deer or lamb*)) and (ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) and (ab=(coliform or ehec or o157) or ti=(coliform or ehec or o157) or ke=(coliform or ehec or o157))

24.
(ab=(sheep or goat* or deer or lamb*) or ti=(sheep or goat* or deer or lamb*) or ke=(sheep or goat* or deer or lamb*)) and (ab=(probiotic* or lactobac* or bifidobac* or

propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) and (ab=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157) or ti=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157) or ke=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157))

26.
(ab=(sheep or goat* or deer or lamb*) or ti=(sheep or goat* or deer or lamb*) or ke=(sheep or goat* or deer or lamb*)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(feces or faeces or fecal or faecal or manure) or ti=(feces or faeces or fecal or faecal or manure) or ke=(feces or faeces or fecal or faecal or manure))

27.
(ab=(sheep or goat* or deer or lamb*) or ti=(sheep or goat* or deer or lamb*) or ke=(sheep or goat* or deer or lamb*)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(gastrointestinal or enteric) or ti=(gastrointestinal or enteric) or ke=(gastrointestinal or enteric))

28.
(ab=(sheep or goat* or deer or lamb*) or ti=(sheep or goat* or deer or lamb*) or ke=(sheep or goat* or deer or lamb*)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(coliform or ehec or o157) or ti=(coliform or ehec or o157) or ke=(coliform or ehec or o157))

29.
(ab=(sheep or goat* or deer or lamb*) or ti=(sheep or goat* or deer or lamb*) or ke=(sheep or goat* or deer or lamb*)) and (ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157) or ti=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157) or ke=(shiga-like pre/1 toxin or escherichia pre/1 coli or o pre/1 157))

30.
(ab=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ti=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid) or ke=(competitive pre/1 exclusion or strain pre/1 mixture or dietary pre/1 supplementation or lactic pre/1 acid)) and (ab=(bacteria pre/1 load or bacterial pre/1 load or bacteria pre/1 counts or bacterial pre/1 counts or bacteria pre/1 log or bacterial pre/1 log or bacteria pre/1 level or bacterial pre/1 level) or ti=(bacteria pre/1 load or bacterial pre/1 load or bacteria pre/1 counts or bacterial pre/1 counts or bacteria pre/1 log or bacterial pre/1 log or bacteria pre/1 level or bacterial pre/1 level) or ke=(bacteria pre/1 load or bacterial pre/1 load or bacteria pre/1 counts or bacterial pre/1 counts or bacteria pre/1 log or bacterial pre/1 log or bacteria pre/1 level or bacterial pre/1 level))

31.
(ab=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ti=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin) or ke=(probiotic* or lactobac* or bifidobac* or propionibac* or saccharomyces or fermentation or yeast or bacteriocin)) and (ab=(bacteria pre/1 load or bacterial pre/1 load or bacteria pre/1 counts or bacterial pre/1 counts or bacteria pre/1 log or bacterial pre/1 log or bacteria pre/1 level or bacterial pre/1 level) or ti=(bacteria pre/1 load or bacterial pre/1 load or bacteria pre/1 counts or bacterial pre/1 counts or bacteria pre/1 log or bacterial pre/1 log or bacteria pre/1 level or bacterial pre/1 level) or ke=(bacteria pre/1 load or bacterial pre/1 load or bacteria pre/1 counts or bacterial pre/1 counts or bacteria pre/1 log or bacterial pre/1 log or bacteria pre/1 level or bacterial pre/1 level))

TOTAL: 4, 957
Total of duplicates; 2, 963 (many from the Agricola strategy)
Total without duplicates (ESR)= 1, 994 abstracts

Appendix 6

Generic letter asking investigators for study information.

Dear Dr. {},

Our research group has recently initiated a systematic review to {REVIEW OBJECTIVE}.

The validity of systematic reviews is based on the identification of all relevant literature, both published and unpublished, that addresses the research question. We have conducted extensive literature searches on this topic and are currently evaluating if any evidence pertaining to this question has been missed or if there is any work that is currently in progress. Our goal is to include the most up-to-date information in this systematic review.

During our work on this project we came across the {NAME OF CONFERENCE OR NAME OF CURRENT FUNDING DATABASE} in which you were an author on an {ABSTRACT / RESEARCH PROJECT} titled: "{INCLUDE TITLE} ". We were unable to locate a peer-reviewed version of this report. We are kindly asking you to inform us if this work has been published or whether it would be possible to obtain a report of the results? The level of detail required for inclusion of research results into the systematic review is a description of the study design, data collection and laboratory methods, a description of the statistical analysis, and the results. This would be equivalent to the materials / methods and results section of a publication or report. The results of the review will be published in an aggregated form, and we can protect the identity of unpublished or proprietary work.

Our research team greatly appreciates your feedback and thanks you in advance for your assistance in providing this information. Please contact me if you have questions or concerns.

Sincerely,

Appendix 7

Domains and elements for randomized controlled trials

Domain	Elements*
Study question	• **Clearly focused and appropriate question**
Study population	• **Description of study population** • **Specific inclusion and exclusion criteria** • Sample size justification
Randomization	• *Adequate approach to sequence generation* • ***Adequate concealment method used*** • *Similarity of groups at baseline*
Blinding	• ***Double-blinding (e.g., of investigators, caregivers, subjects, assessors, and other key study personnel as appropriate) to treatment allocation***
Interventions	• **Intervention(s) clearly detailed for all study groups (e.g., dose, route, timing for drugs, and details sufficient for assessment and reproducibility for other types of interventions)** • Compliance with intervention • Equal treatment of groups except for intervention
Outcomes	• **Primary and secondary outcome measures specified** • Assessment method standard, valid, and reliable
Statistical Analysis	• ***Appropriate analytical techniques that address study withdrawals, loss to follow-up, missing data, and intention to treat*** • Power calculation • Assessment confounding • Assessment of heterogeneity, if applicable
Results	• **Measure of effect for outcomes and appropriate measure of precision** • Proportion of eligible subjects recruited into study and followed up at each assessment
Discussion	• **Conclusions supported by results with possible biases and limitations taken into consideration**
Funding or sponsorship	• ***Type and source of support for study***

*Elements appearing in italics are those with an empirical basis. Elements appearing in bold are those considered essential to give a system a full Yes rating for the domain and thereby pass the quality assessment stage.

From AHRQ (2002).

Domains and elements for observational studies

Domains	Elements
Study Question	• **Clearly focused and appropriate question**
Study Population	• **Description of study populations** • Sample size justification
Comparability of subjects+	For all observational studies • Specific inclusion/exclusion criteria for all groups • Criteria applied equally to all groups • Comparability of groups at baseline with regard to disease status and prognostic factors • Study groups comparable to non-participants with regard to confounding factors • *Use of concurrent controls* • Comparability of follow-up among groups at each assessment Additional criteria for case-control studies • Explicit case definition • Case ascertainment not influenced by exposure status • Controls similar to cases except without condition of interest and with equal opportunity for exposure
Exposure or intervention	• **Clear definition of exposure** • Measurement method standard, valid and reliable • Exposure measured equally in all study groups
Outcome measurement	• **Primary/secondary outcomes clearly defined** • Outcomes assessed blind to exposure or intervention status • Method of outcome assessment standard, valid and reliable • Length of follow-up adequate for question
Statistical Analysis	• Statistical tests appropriate • Multiple comparisons taken into consideration • Modeling and multivariate techniques appropriate • Power calculation provided • **Assessment of confounding** • Does response assessment, if appropriate
Results	• **Measure of effect for outcomes and appropriate measure of precision** • Adequacy of follow-up for each study group
Discussion	• **Conclusions supported by results with biases and limitations taken into consideration**
Funding or Sponsorship	• ***Type and sources of support for study***

*Elements appearing in italics are those with an empirical basis. Elements appearing in bold are those considered essential to give a system a Yes rating for the domain.

+Domain for which a Yes rating required that a majority of elements be considered.

From report prepared by: Research Triangle Institute-University of North Carolina Evidence-based Practice Center to the Agency for Health Research and Quality

APPENDIX 8

References for critical review of the literature

1. Dohoo, IR, Waltner-Toews, D: Interpreting clinical research. Part I. General considerations. Compend Contin Educ Pract Vet 7:S473-S478, 1985.

2. Dohoo, IR, Waltner-Toews, D: Interpreting clinical research. Part II. Descriptive and experimental studies. Compend Contin Educ Pract Vet 7:S513-S519, 1985.

3. Dohoo, IR, Waltner-Toews, D: Interpreting clinical research. Part III. Observational studies and interpretation of results. Compend Contin Educ Pract Vet 7:S605-S613, 1985.

4. Martin, SW, Bonnett BN: Clinical epidemiology. Can Vet J 28:318-325, 1987.

5. Evidence-Based Medicine Working Group: Evidence-based medicine: A new approach to teaching the practice of medicine. J Am Med Assoc 268:2420-2425, 1992.

6. Guyatt, GH, Sackett, MD, Cook, DJ (for the Evidence-Based Medicine Working Group): Users' Guides to the Medical Literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? J Am Med Assoc 270:2598-2601, 1993.

7. Guyatt, GH, Sackett, MD, Cook, DJ (for the Evidence-Based Medicine Working Group): Users' Guides to the Medical Literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patient? J Am Med Assoc 271:59-63, 1994.

8. Jaeschke, R, Guyatt, GH, Sackett, MD (for the Evidence-Based Medicine Working Group): Users' Guides to the Medical Literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? J Am Med Assoc 271:389-391, 1994.

9. Jaeschke, R, Guyatt, GH, Sackett, MD (for the Evidence-Based Medicine Working Group): Users' Guides to the Medical Literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? J Am Med Assoc 271:703-707, 1994.

10. Levine, M, Walter, S, Lee H, et al (for the Evidence-Based Medicine Working Group): Users' Guides to the Medical Literature. IV. How to use an article about harm. J Am Med Assoc 271:1615-1619, 1994.

11. Laupacis, A, Wells, G, Richardson, S, et al (for the Evidence-Based Medicine Working Group): Users' Guides to the Medical Literature. V. How to use an article about prognosis. J Am Med Assoc 272:234-237, 1994.

12. Oxman, AD, Cook, DJ, Guyatt, GH (for the Evidence-Based Medicine Working Group): Users' Guides to the Medical Literature. VI. How to use an overview. J Am Med Assoc 272:1367-1371, 1994.

13. Rennie, D, Flanagin, A: Publication bias: The triumph of hope over experience. J Am Med Assoc 267:411-412, 1992.

14. Cochrane, AL: 1931-1971: A critical review, with particular reference to the medical profession. *In*: Medicine for the Year 2000. London, Office of Health Economics, p 2-12. 1979.

15. Curtis, CR, Salman, MD and Shott, SP. Values. JAVMA. Aug 1; 197(3): 318-320, 1990.

16. Curtis, CR Salman, MD and Shott, S. Power and Sample Size. JAVMA. Oct 1; 197(1): 838-840, 1990.

17. Salman, MD, Curtis, CR and Shott, S. Comparing Means. JAVMA. Jan 1; 198(1): 62-65, 1991.

18. Salman, MD, Curtis, CR and Shott, S. Data Description. JAVMA. Jul 1; 197(1): 36-38, 1990.

19. Shott, S. Association. JAVMA. Feb 1; 198(3): 404-407, 1991.

20. Shott, S. Comparing Proportions. JAVMA. Dec 1; 197(11): 1460-1462, 1990.

21. Shott, S. Confidence Intervals. JAVMA. Sep 1; 197(5): 576-578, 1990.

22. Shott, S. Logistic Regression and Discriminant Analysis. JAVMA. Jun 1; 198(11): 1902-1904, 1991.

23. Shott, S. Nonparametric Statistics. JAVMA. Apr 1; 1988(7): 1126-1128, 1991.

24. Shott, S. Regression. JAVMA.Mar 1; 198(5): 798-801, 1991.

25. Shott, S. Study Design. JAVMA. Nov 1; 197(9): 1142-1144, 1990.

26. Shott, S. Survival Analysis. JAVMA.May 1; 198(9): 1513-1515, 1991.

FORM 1

RELEVANCE TOOL FOR SCREENING ABSTRACTS, GENERIC

(modified from PHRED documents [Available at: http://www.phred-redsp.on.ca/])

Ref ID_____
Reviewer:_____

Relevance Criteria

1) Does the abstract describe a primary research study (as opposed to a review) Y N

2) Does the abstract describe the intervention addressed in the systematic review question in the species of interest Y N

3) Does the abstract assess the intervention in relation to the Outcome(s) of interest for the systematic review question Y N

Must answer "yes" to all questions for the reference to advance to the quality assessment stage.

FORM 2
RELEVANCE TOOL FOR SCREENING ABSTRACTS,
"Probiotics for the reduction of *E. coli* O157 in the feces of beef and dairy cattle"

Ref ID_____
Reviewer:_____

Relevance Criteria

1) Does this abstract describe primary research (as opposed to a review)?
Y N

2) Does the abstract describe the use of probiotics in live post-weaned domestic ruminants? Y N

3) Does the abstract describe the effect of probiotics on the presence or level of *E. coli* O157 in faeces? Y N

FORM 3
RELEVANCE TOOL FOR SCREENING ABSTRACTS,
Multiple questions in one review (Johne's disease / Crohn's disease example)

1) Does this abstract describe primary research (as opposed to a review)?
Y N

2) Does this abstract investigate (check all that apply)?

- A potential association between Crohn's disease in humans and *Mycobacterium avium subspecies paratuberculosis* (MAP) isolated from humans?
- A potential association between Johne's disease (MAP or paratuberculosis) in ruminants and Crohn's disease in humans?
- Dairy products or human food as a potential source of MAP?

\

FORM 4
A. CHECKLIST FOR APPRAISING THE QUALITY OF RANDOMIZED CLINICAL TRIALS

Quality item	Coding	Explanation
Objectives and Study Population		
Do the objectives address the systematic review question?	Yes No	Yes: Objectives clearly stated. No: Return to relevance screening stage and re-evaluate.
Was the sample size justified?	Yes Partial No	Yes: Use of sample-size formulas, based on desired power or precision and estimate of expected variability to detect differences. Partial: Informal guesses of a sample size. No: No details in the text.
Were the animals housed or grouped in a way that is representative of field conditions?	Yes Partial No	Yes: Animals housed in densities, pens and rooms representative of field conditions. Partial: Animals housed or grouped in small densities, not similar of field conditions. No: Animals housed or grouped individually.
Was the reason and proportion of livestock operations that declined participation described?	Yes No Single farm	Yes No Single farm study.
Intervention (treatment allocation, blinding)		
Were sampling units randomly assigned to the treatment groups?	Yes Partial No	Yes: computer or random numbers table, *a-priori* assignment of tagged numbers, alternation or systematic allocation, stratified random sample, cluster randomization. Partial: 'randomized' or randomly allocated without explanation, a day assignment. No: Sample drawn without a formal process of random selection: judgment, convenience, purposive.
Prior to the intervention, were the sampling units tested for the outcome disease?	Yes No	Yes No
Were the intervention protocols adequately described?	Yes No	Yes No
Were the route, administration schedule, and animal grouping level of interventions feasible in a commercial setting?	Yes No	Yes No
Was an appropriate control group used?	Yes No	Yes No
Was the outcome assessor appropriately blinded to the intervention status of the treatment units?	Yes No Unknown	Yes No Unknown

Quality item	Coding	Explanation
Outcome		
Were laboratory tests used to determine the outcome described and adequate?	Yes No	Yes: * No * Minimum standards have to be developed by reviewer experts for each systematic review.
Was the time from intervention administration to measurement of outcome sufficient to have the outcome of interest?	Yes No	Yes: Study allows enough time to observe the outcome of interest after the intervention is performed. No * Define for each systematic review.
Withdrawals and loss to follow-up		
Were mortality, withdrawals and/or loss to follow-up reported?	Yes Partial No	Yes: Numbers stated or deducible from tables and reasons provided for each group or no losses. Partial: numbers but not reasons (or vice versa). No
Was the proportion of lost to follow-up adequate?	Yes No	Yes: Percentages of lost of subjects <10%. No: >10% or not described.
Data analysis		
Was the statistical analysis appropriate?	Yes No	Yes: analysis fits design, appropriate analysis of clustered data when required. No
Were the estimates and measures of variability used to address the research question presented adequately?	Yes No	Yes: parameter estimates + measure of variability or P value provided or sufficient data provided for post-hoc corrected statistics. No
Were confounders appropriately considered?	Yes No	Yes: include exclusion, matching or analytical control. Partial: some confounders controlled but no all of those identified as important. No * Important confounders defined for each systematic review
Conclusions		
Were conclusions supported by the results?	Yes No	Yes No

Adapted from CRD and AHRQ manuals

Note: Elements appearing in bold are those considered essential to give a system a final "yes" rating.
Abstract assigned a "no" to any bold question will be excluded from the review.

B. CHECKLIST FOR APPRAISING THE QUALITY OF CHALLENGE TRIALS

Quality item	Coding	Explanation
Objectives and Study Population		
Do the objectives address the systematic review question?	Yes No	Yes: Objectives clearly. No: Return to relevance screening stage and re-evaluate.
Was the sample size justified?	Yes Partial No	Yes Use of sample-size formulas, based on desired power or precision and estimate of expected variability to detect differences. Partial: Informal guesses of a sample size. No: No details in the text.
Intervention (challenge administration, treatment allocation, blinding)		
Were sampling units randomly assigned to the treatment groups?	Yes Partial No	Yes: computer or random numbers table, *a-priori* assignment of tagged numbers, alternation or systematic allocation, stratified random sample, cluster randomization. Partial: 'randomized' or randomly allocated without explanation, a day assignment. No: Sample drawn without a formal process of random selection: judgment, convenience, purposive.
Prior to the intervention, were the sampling units tested for the outcome disease?	Yes No	Yes No
Were the intervention protocols adequately described?	Yes No	Yes No
Were the route, administration schedule, and animal grouping level of interventions feasible in a commercial setting?	Yes No	Yes No
Was an appropriate control group used?	Yes No	Yes No
Was the outcome assessor appropriately blinded to the intervention status of the treatment units?	Yes No Unknown	Yes No Unknown
Outcome assessment		
Were laboratory tests to determine the outcome described and adequate?	Yes No	Yes: * No * Minimum standards have to be developed by re viewer experts for each systematic review.
Was the time from intervention administration to measurement of outcome sufficient to determine the outcome of interest?	Yes No	Yes: Study allows enough time to observe the outcome of interest after the intervention is performed. No *Define for each systematic review.

Quality item	Coding	Explanation
Withdrawals and loss to follow-up		
Were mortality, withdrawals and/or loss to follow-up reported?	Yes Partial No	Yes: Numbers stated or deducible from tables and reasons provided for each group or no losses. Partial: numbers but not reasons (or vice versa). No
Was the proportion of lost to follow-up adequate?	Yes No	Yes: Percentages of lost of subjects <10% No: >10% or not described.
Data analysis		
Was the statistical analysis appropriate?	Yes No	Yes: analysis fits design, appropriate analysis of clustered date when required. No
Were the estimates and measures of variability used to address the research question presented adequately?	Yes No	Yes: parameter estimates + measure of variability or P value provided or sufficient data provided for post-hoc corrected statistics. No
Were confounders appropriately considered?	Yes No	Yes: include exclusion, matching or analytical control. Partial: some confounders controlled, but not all of those identified as important. No * Important confounders defined for each systematic review.
Conclusions		
Were conclusions supported by the results?	Yes No	Yes: No:

Adapted from CRD and AHRQ manuals

Note: Elements appearing in bold are those considered essential to give a system a final "yes" rating.
Abstract assigned a "no" to any bold question will be excluded from the review.

C. CHECKLIST FOR APPRAISING THE QUALITY OF COHORT STUDIES

Quality item	Coding	Explanation
Objectives and Study Population		
Do the objectives address the systematic review question?	Yes No	Yes: Objectives clearly stated. No: Return to relevance screening stage and re-evaluate.
Was the sample size justified?	Yes Partial No	Yes Use of sample-size formulas, based on desired power or precision and estimate of expected variability to detect differences. Partially: Informal guesses of a sample size. No: No details in the text.
Were the animals housed or grouped in a way that is representative of field conditions?	Yes Partial No	Yes: Animals housed in densities, pens and rooms representative of field conditions. Partial: Animals housed or grouped in small densities, not similar to field conditions. No: Animals housed or grouped individually.
Was the reason and proportion of livestock operations that declined participation described?	Yes No Single farm	Yes No Single farm study
Intervention (treatment allocation, blinding)		
Within the operations, was animal or pen selection described and justified?	Yes No	Yes No
Were the route, administration schedule, and animal grouping level of interventions feasible in a commercial setting?	Yes No	Yes: No
Was and appropriate control group used?	Yes No	Yes No
Was the outcome assessment appropriately blinded to the intervention status of the treatment units?	Yes No Unknown	Yes: No Unknown
Outcome assessment		
Were laboratory tests to determine the outcome described and adequate?	Yes No	Yes: * No * Minimum standards have to be developed by reviewer experts for each Systematic review.
Was the time from treatment administration to measurement of outcome sufficient to have the outcome of interest?	Yes No	Yes: Study allows enough time to observe the outcome of interest after the intervention is performed. No * Define for each systematic review.

Quality item	Coding	Explanation
Withdrawals and loss to follow-up		
Were mortality, withdrawals and/or loss to follow-up reported?	Yes Partial No	Yes: Numbers stated or deducible from tables and reasons provided for each group or no losses. Partial: numbers but not reasons (or vice versa). No
Was the proportion of lost to follow-up adequate?	Yes No	Yes: Percentages of lost of subjects <10% No: >10% or not described
Data analysis		
Was the statistical analysis appropriate?	Yes No	Yes: analysis fits design, appropriate analysis of clustered date when required. No
Were the estimates and measures of variability used to address the research question presented adequately?	Yes No	Yes: parameter estimates + measure of variability and/or P value provided or sufficient data provided for post-hoc corrected statistics. No
Were confounders appropriately considered?	Yes No	Yes: includes exclusion, matching or analytical control. Partial: some confounders controlled but not all of those identified as important. No * Important confounders defined for each systematic review.
Conclusions		
Were conclusions supported by the results?	Yes No	Yes: No:

Adapted from CRD and AHRQ manuals

Note: Elements appearing in bold are those considered essential to give a system a final "yes" rating.
Abstract assigned a "no" to any bold question will be excluded from the review.

D. CHECKLIST FOR APPRAISING THE QUALITY OF CASE-CONTROL AND CROSS-SECTIONAL STUDIES

Quality item	Coding	Explanation
Objectives and Study Population		
Do the objectives address the systematic review question?	Yes No	Yes: Objectives clearly. No: Return to relevance screening stage and re-evaluate.
Was the sample size justified?	Yes Partial No	Yes Use of sample-size formulas, based on desired power or precision and estimate of expected variability to detect differences. Partially: Informal guesses of a sample size. No: No details in the text.
Were the animals housed or grouped in a way that is representative of field conditions?	Yes Partial No	Yes: animals housed in densities, pens and rooms representative of field conditions. Partial: animals housed or grouped in small densities, not similar of field conditions. No: animals housed or grouped individually.
Was the reason and proportion of livestock operations that declined participation described?	Yes No Single farm	Yes No Single farm study.
Were cases and controls similar, except for the condition of interest but with equal opportunity for exposure? **(Only case-control studies)**	Yes No	Yes: controls selected from the same study base population and remained free of the outcome during the study period. No Not applicable: for cross-sectional studies.
Intervention (treatment allocation, blinding)		
Within the operations, was animal or pen selection described and justified?	Yes No	Yes No
Were the intervention protocols or exposure variable adequately described?	Yes No	Yes No
Was an appropriate comparison group for the intervention or exposure variable used?	Yes No	Yes No
Was the intervention or exposure variable assessed equally for cases and controls?	Yes No	Yes No
Outcome assessment		
Were laboratory tests to determine the outcome described and adequate?	Yes No	Yes: * No * Minimum standards have to be developed by re viewer experts for each systematic review.
Was the time from treatment administration to measurement of outcome sufficient to have the outcome of interest?	Yes No	Yes: Study allows enough time to observe the outcome of interest after the intervention is performed. No *Define for each systematic review.

Quality item	Coding	Explanation
Withdrawals and loss to follow-up		
Were mortality, withdrawals and/or loss to follow-up reported?	Yes Partial No	Yes: Numbers stated or deducible from tables and reasons provided for each group or no losses. Partial: numbers but not reasons (or vice versa). No
Was the proportion of lost to follow-up adequate?	Yes No	Yes: Percentages of lost of subjects <10% No: >10% or not described
Data analysis		
Was the statistical analysis appropriate?	Yes No	Yes: analysis fits design, appropriate analysis of clustered date when required. No
Were the estimates and measures of variability used to address the research question presented adequately?	Yes No	Yes: parameter estimates + measure of variability and/or P value provided or sufficient data provided for post-hoc corrected statistics. No
Were confounders appropriately considered?	Yes No	Yes: includes exclusion, matching or analytical control. Partial: some confounders controlled but not all of those identify as important. No * Important confounders defined for each systematic review.
Conclusions		
Were conclusions supported by the results?	Yes No	Yes No

Adapted from CRD and AHRQ manuals

Note: Elements appearing in bold are those considered essential to give a system a final "yes" rating.
Abstract assigned a "no" to any bold question will be excluded from the review.

FORM 5
DATA EXTRACTION FORM FOR EXAMPLE REVIEW ON THE USE OF PROBIOTICS FOR REDUCING *E. COLI* O157 IN RUMINANTS.

Data Extraction Form		
General Information	**ESR**	**Explanation**
Ref ID	Included in ESR	
Author (s)	Included in ESR	
Article Title, volume, pages	Included in ESR	
Source	Included in ESR	Journal, Proceedings etc.
Date of data extraction	Included in ESR	T-box (yyyy/mm/dd) *If revisions are necessary*
Name of person that performed data extraction	Included in ESR	
Language of publication	English Other language T-box	*Including papers in another language decreases bias*
Funding of the study	T-box (name of Institution) Unknown	
Study Design, check one:	a. Randomized controlled trial b. Challenge study c. Case-control d. Cross sectional e. Cohort study f. Prevalence study	*This question indicates the level of evidence provided by the study*
Population		
Country where study subjects were located	USA Canada Europe T-box South-America T-box Other T-box	Enter all that apply in alphabetical order *External relevance, heterogeneity*
Place where the study was performed?	a. Single farm- Commercial b. Single farm – Experimental station or University T-box c. More than one farm T-box (number of farms) d. Not described e. Other (T-box)	T-box (Name of experimental station or University, number of farms) *External validity*
What type of sampling was used to select the farm(s)? **(Observational Studies)**	a. Convenience T-box b. Purposive T-box c. Random sample T-box d. Not described e. Not applicable (RTC, Challenge trials)	T-box explain why convenience or type of random sample *External and internal validity*

What were the inclusion/ exclusion criteria of the experimental units? **(Observational Studies)**	a. Explain T-box b. Not described c. Not applicable	***T-box E.g. cull, lactating, dry and cows in sick pens to represent the distribution of cows within herd in terms of parity and classification*** *External and internal validity*
Attributes of animals	a. Commercial livestock or equivalent b. Colostrums deprived c. Gnotobiotic d. Deliberately immuno-suppressed e. Not described f. Other	*External validity*
Which type of animals where used in the experiment? Check all that apply	a. Dairy cows b. Beef (housed in pens) c. Beef (on pasture) d. Sheep e. Goats f. Other T-box	*External validity*
What was the age of the experimental units?	a. Adult b. Weaning juveniles c. Pre-weaned d. Neonates e. Mixed groups f. Other T-box g. Not described	*External validity*
What was the breed of the animals used in the study? Check all that apply	a. Dairy b. British X c. Continental X d. Brahman X e. Mixed f. Suffolk g. Other T-box h. Not described	*External validity*
What was the gender of the animals used in the study?	a. Females b. Males c. Both d. Not described	*External validity*
Intervention		
What was the experimental unit level at which treatment was allocated?	a. Farm b. Pen c. Individual animal housed alone d. Individual animal within pens (all animals are experimental) e. Individual animal housed within pens (that contain also not experimental animals) f. Not described	*External validity and study methodology*

What type of allocation to the treatment groups was used? **(Randomized and Challenge Trials)**	a. Simple randomization T-box b. Stratified randomization T-box c. Blocked randomization T-box d. Systematic assignment T-box e. Not described f. Not applicable	T box (describe method for b, c, and d) *External and internal validity*
Were the animals negative to *E. coli* O157 prior to the start of the experiment?	Yes No Not tested Not applicable	Only for RCT, Challenge trials and Cohorts *Descriptive purposes*
Were the animals challenged with *E. coli* before, during or after the administration of the probiotic?	Before During After Not described Not applicable	Only for RCT, Challenge trials and Cohorts *Descriptive purposes*
What was the level of *E. coli* challenge?	T-box	Include number, level and units e.g. 5*10^8 CFU/g *Descriptive purposes*
What type of probiotic strain given to each of the treatment groups, the dose and units? Group A Fill in appropriate information	a. *Saccharomyces cervisiae* T-box b. *Streptococcus thermophilus* T-box *c. Enterococcus faecium* T-box *d. Lactobacillus acidophilus* T-box e. *Lactobacillus casei* T-box f. *Lactobacillus fermentum* T-box g. *Propionibacterium freudenreichii* T-box h. *E. coli* probiotics T-box i. Only Direct Fed Microbial or Probiotic	T-box (strain, dose, unit) e.g NP45, 5*10^6 CFU/ml *Descriptive purposes*
Group B Fill in appropriate information	a. *Saccharomyces cervisiae* T-box b. *Streptococcus thermophilus* T-box *e. Enterococcus faecium* T-box *f. Lactobacillus acidophilus* T-box g. *Lactobacillus casei* T-box h. *Lactobacillus fermentum* T-box i. *Lactobacillus plantrum* T-box j. *Propionibacterium freudenreichii* T-box k. *Clostridium butyricum* T-box h. *E. coli* probiotics T-box i. Direct Fed Microbial or Probiotic	Same
Group C	Same	Same
What type of non-probiotic control group(s) was used?	a. Not infected not probiotics b. Infected not probiotics c. Not infected not probiotic but placebo. d. Infected not probioitc but placebo. e. Not described f. Not applicable	*Descriptive purposes* *Internal validity*

What was the length of administration of the probiotic?	T-box	T-box (total time of administration of probiotic determine number and unit of time (e.g 3 weeks) *Descriptive purposes*
What was the frequency of administration of the probiotic?	T-box	T-box (determine number and unit of time (e.g. 2x/d or 1x/at the beginning) *Descriptive purposes*
How were probiotics administered?	a. Orally through the diet T-box b. Orally through the water T-box c. Orally through milk T-box d. Rumen cannula T-box e. Other T-box e. Not described	*Descriptive purposes*
Outcome		
What was the technique used to measure the presence of *E. coli* O157? Check all that apply and explain	a. Direct plating T-box b. Culture with enrichment medium T-box c. Culture with selective medium T-box d. Immunomagnetic separation e.Agglutination test f. Serological test T-box g. Biochemical test T-box h. Molecular technique T-box i. Dilutions for counting of bacteria T-box j. Other T-box	Direct plating (McConckey) Enrichment (specify enrichment medium used) Selective (specify selective medium used) Biochemical test (API system) Molecular technique (PCR) Dilution for counting (10 fold dilution in PBS and direct plating for counting) *Descriptive purposes*
Were +/- controls, or systematic sample of colonies used for the tests used above? Check all that apply and explain	a. Direct plating T-box b. Culture with enrichment medium T-box c. Culture with selective medium T-box d. Immunomagnetic separation e. Serological test T-box f. Biochemical test T-box g. Molecular technique T-box h. Dilutions for counting of bacteria T-box i. Other T-box j. No controls described T-box	T-box explain what type of positive or negative control were used or number of aliquots used in dilutions *Validity and methodology of laboratory tests performed*
Which secondary outcomes measured to determine the effect of the probiotics? Check all that apply	a. Average daily gain ADG b. Average daily feed intake ADFI c. Feed:Gain ratio F:G d. Scores of diarrhea e. Carcass evaluation f. Other T-box	*Descriptive purposes – beneficial and harmful effects of intervention*

	g. None	
Was blinding of people administering the intervention and evaluating secondary outcomes reported?	Yes T-box No Not described Not applicable	T-box specify in cases where scores of diarrhea are evaluated **(if we include this information)** *Internal validity*
Statistical Analysis		
How many farms, pens and animals were included in the experiment? Check all that apply	Farm T-box Blocks T-box Pens T-box Animals T-box Not described	CHECK ALL THAT APPLY AND INDICATE THE NUMBER *How analysis of data was performed*
How many farms, pens and animals were included in each treatment group? Check all that apply	Farm T-box Blocks T-box Pens T-box Animals T-box Not described	
Was the statistical analysis performed at the farm level?	Yes No Not described Not applicable	*Control of cluster data*
Was the statistical analysis performed at the pen level?	Yes No Not described Not applicable	*Control of cluster data*
If statistical analysis was performed at pen level, was farm level controlled for?	Yes No Not described Not applicable	*Control of cluster data*
Was the statistical analysis performed at an individual level?	Yes No Not described Not applicable	*Control of cluster data*
If statistical analysis was performed at an individual level, was farm level controlled for?	Yes No Not described Not applicable	*Control of cluster data*
If statistical analysis was performed at an individual level, was pen level controlled for?	Yes No Not described Not applicable	*Control of cluster data*
Was the outcome of interest measured more than once?	Yes T-box No Not described	T-box (explain frequency eg 2x/15 of trial, every day) *Repeated measures*
If the outcome of interest was measured more than once, was time controlled in the analysis?	Yes T-box No Not described	T-box (indicate estimate, CI and/or P value of time effect)

	Not applicable	*Repeated measures*
Was raw data of outcome of interest presented?	Yes No	If yes, then we have to go back to the paper and analyzed the data.
Were confounders age, sex, weight controlled by blocking, stratifying, matching or controlled within analysis? Check all that apply and explain how they were controlled	Age T-box Sex T-box Weight T-box Other T-box Not described T-box	*Confounding*
When samples were collected? Check all that apply	Spring Summer Fall Winter	
Was any effect of season reported? If yes, explain	Yes T-box No Not reported Not applicable	More for Observational studies T-box (explain the effect of season reported)
Results		
Table of Results-Continuous Data **Fill in the information given** Group A Group B Group C Group D Control	- OVERALL, LEAST SQUARE OR CONTRAST MEANS - Log mean differences from control - Unit of results - Lower/ upper 95% CI - SE - P value	Set as table in ESR *How continuous results reported*
Table of Results – Dichotomous Data **Fill in the information given** Group A Group B Group C Group D Control	- No. of + experimental units/group - Proportion of + experimental units/group - Total No. of experimental units/group - OR - RR - Lower/Upper 95% CI - P value	Set as table in ESR *How continuous results reported*
Tables of secondary outcomes **Fill in the information given**		

Average Daily Gain (ADG) Group A Group B Group C Group D Control	- Overall, least square or contrast Means - Mean differences from control - Unit of result - Lower/ upper 95% CI - SE - P value	Set as a table in ESR Not applicable for all studies
Average Feed Intake (ADFI) Group A Group B Group C Group D Control	- Overall, least square or contrast Means - Mean differences from control - Unit of result - Lower/ upper 95% CI - SE - P value	Set as a table in ESR Not applicable for all studies
Feed:Gain (F:G) Group A Group B Group C Group D Control	- Overall, least square or contrast Means - Mean differences from control - Unit of result - Lower/ upper 95% CI - SE - P value	Set as a table in ESR Not applicable for all studies
Scores of diarrhea Group A Group B Group C Group D Control	- Overall, least square or contrast Means - Mean differences from control - Unit of result - Lower/ upper 95% CI - SE - P value	Set as a table in ESR Not applicable for all studies
Final Comment		
Comments/Concerns	Text-box	T-box (*indicate any problems and/or additional information that might be important for summarizing the data*)