2007

Evaluation of cell mediated immune diagnostic tests to detect Mycobacterium avium subspecies paratuberculosis

Suelee Robbe-Austerman

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Agriculture Commons, Animal Sciences Commons, Large or Food Animal and Equine Medicine Commons, and the Veterinary Pathology and Pathobiology Commons

Recommended Citation


https://lib.dr.iastate.edu/rtd/15970

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Evaluation of cell mediated immune diagnostic tests to detect *Mycobacterium avium* subspecies *paratuberculosis*

by

Suelee Robbe-Austerman

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Veterinary Microbiology

Program of Study Committee:
Ronald Griffith, Major Professor
John Bannantine
Rich Evans
Annette O’Connor
Judith Stabel

Iowa State University
Ames, Iowa
2007
# TABLE OF CONTENTS

ABSTRACT.................................................................................................................. v

CHAPTER 1. INTRODUCTION....................................................................................... 1
  1.1 Dissertation organization.................................................................................. 1
  1.2 Research summary.......................................................................................... 2
  1.3 References.................................................................................................... 4

CHAPTER 2. REVIEW OF LITERATURE........................................................................ 5
  2.1 The discovery of paratuberculosis and development of johnin....................... 5
  2.2 The use of johnin and avian tuberculin intravenously in the United States to diagnose paratuberculosis................................................................. 7
  2.3 Development of the intradermal skin test for paratuberculosis......................... 8
  2.4 Improving the growth and production of johnin............................................. 9
  2.5 The consequences of repeated skin testing with johnin................................. 12
  2.6 Comparison of the skin test and the complement fixation test......................... 12
  2.7 Field utilization of the skin test in tuberculosis vs. paratuberculosis................. 15
  2.8 Field trials correlating evidence of paratuberculosis with skin test positivity in cattle........................................................................................................... 15
  2.9 Using the skin test in small ruminants to diagnose paratuberculosis............... 18
  2.10 Fecal culture as a diagnostic test for paratuberculosis................................... 19
  2.11 The lymphocyte transformation test............................................................. 20
  2.12 The development of the gamma Interferon ELISA........................................ 20
  2.13 Impact of the gamma interferon ELISA in diagnosing tuberculosis............. 22
  2.14 Field trials correlating evidence of paratuberculosis with the gamma interferon ELISA in cattle............................................................... 23
  2.15 Using the gamma interferon ELISA in small ruminants to diagnose paratuberculosis..................................................................................................... 25
  2.16 Research objectives....................................................................................... 26
  2.17 References................................................................................................... 27

CHAPTER 3. SENSITIVITY AND SPECIFICITY OF THE AGAR-GEL-IMMUNODIFFUSION TEST, ELISA AND THE SKIN TEST FOR DETECTION OF PARATUBERCULOSIS IN UNITED STATES MIDWEST SHEEP POPULATIONS................................................................. 37
  3.1 Abstract......................................................................................................... 38
  3.2 Introduction.................................................................................................... 39
  3.3 Materials and methods................................................................................... 41
    3.3.1 Non-infected sheep population................................................................. 41
    3.3.2 Infected sheep population....................................................................... 42
    3.3.3 Skin testing and serum collection procedure......................................... 43
    3.3.4 AGID testing............................................................................................ 43
    3.3.5 ELISA testing.......................................................................................... 44
    3.3.6 Necropsy................................................................................................. 44
CHAPTER 4. EVALUATION OF THE GAMMA INTERFERON ELISA IN SHEEP SUBCLINICALLY INFECTED WITH MYCOBACTERIUM AVIUM SUBSPECIES PARATUBERCULOSIS USING A WHOLE CELL SONICATE OR A JOHNIN PURIFIED PROTEIN DERIVATIVE

CHAPTER 5. TIME DELAY, TEMPERATURE EFFECTS AND ASSESSMENT OF POSITIVE CONTROLS ON WHOLE BLOOD FOR THE GAMMA INTERFERON ELISA TO DETECT PARATUBERCULOSIS
### 5.7 References

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.7 References</td>
<td>90</td>
</tr>
</tbody>
</table>

### CHAPTER 6. SKIN TEST AND GAMMA INTERFERON ENZYME-LINKED IMMUNOSORBENT ASSAY RESULTS IN SHEEP EXPOSED TO DEAD *MYCOBACTERIUM AVIUM* SUBSPECIES *PARATUBERCULOSIS* ORGANISMS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Abstract</td>
<td>98</td>
</tr>
<tr>
<td>6.2 Introduction</td>
<td>99</td>
</tr>
<tr>
<td>5.3 Materials and methods</td>
<td>100</td>
</tr>
<tr>
<td>5.4 Results</td>
<td>102</td>
</tr>
<tr>
<td>5.5 Discussion</td>
<td>103</td>
</tr>
<tr>
<td>5.6 Sources and manufacturers</td>
<td>104</td>
</tr>
<tr>
<td>5.7 References</td>
<td>105</td>
</tr>
</tbody>
</table>

### CHAPTER 7 GENERAL CONCLUSIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1 General Discussion</td>
<td>107</td>
</tr>
<tr>
<td>7.2 Comparison of the IFN-γ ELISA and the skin</td>
<td>110</td>
</tr>
<tr>
<td>7.3 Weakness of the studies in this thesis</td>
<td>111</td>
</tr>
<tr>
<td>7.4 Preliminary Results of a longitudinal study evaluating the skin test and IFN-γ ELISA in calves</td>
<td>114</td>
</tr>
<tr>
<td>7.5 Recommendations for future research</td>
<td>115</td>
</tr>
</tbody>
</table>

### APPENDIX A. SPECIFICITY DIFFERENCES IN ANTIGENS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPENDIX A. SPECIFICITY DIFFERENCES IN ANTIGENS</td>
<td>117</td>
</tr>
</tbody>
</table>

### APPENDIX B. REPEATABILITY OF THE SKIN TEST

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPENDIX B. REPEATABILITY OF THE SKIN TEST</td>
<td>118</td>
</tr>
</tbody>
</table>

### APPENDIX C. FALSE POSITIVE IFN-γ ELISA RESULTS IN A SHEEP FLOCK

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPENDIX C. FALSE POSITIVE IFN-γ ELISA RESULTS IN A SHEEP FLOCK</td>
<td>119</td>
</tr>
</tbody>
</table>

### APPENDIX D. ALTERNATIVE RESULTS USING THE TAGS PROGRAM TO ESTIMATE SENSITIVITY AND SPECIFICITY WITHOUT A GOLD STANDARD

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPENDIX D. ALTERNATIVE RESULTS USING THE TAGS PROGRAM TO ESTIMATE SENSITIVITY AND SPECIFICITY WITHOUT A GOLD STANDARD</td>
<td>120</td>
</tr>
</tbody>
</table>

### APPENDIX E. PRELIMINARY RESULTS OF THE SKIN TEST AND IFN-γ ELISA FROM CALVES BORN TO MAP INFECTED DAMS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPENDIX E. PRELIMINARY RESULTS OF THE SKIN TEST AND IFN-γ ELISA FROM CALVES BORN TO MAP INFECTED DAMS</td>
<td>121</td>
</tr>
</tbody>
</table>

### ACKNOWLEDGEMENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>123</td>
</tr>
</tbody>
</table>
ABSTRACT

*Mycobacterium avium* subspecies *paratuberculosis* (MAP) causes Johne’s disease, a chronic progressive fatal disease of ruminants. Similar to tuberculosis, diagnostic tests based on detecting the organism or antibodies are not capable of detecting infected animals in the early stages. Consequently, diagnostic tests based upon the cell mediated immune response, such as the skin test and gamma interferon enzyme-linked immunoblot assay (ELISA), which have been successful at detecting tuberculosis need to be investigated further for paratuberculosis. The first paper evaluated the skin test in sheep along with two antibody based diagnostic tests using Bayesian statistical methods that estimate the sensitivity and specificity of diagnostic tests in multiple populations without a gold standard. The second paper used receiver operating characteristic (ROC) analysis to evaluate the accuracy of the gamma interferon (IFN-γ) ELISA in subclinically infected sheep. These 2 studies similarly concluded the skin test and IFN-γ ELISA had an estimated sensitivity of around 70% and a specificity of 98%. The next paper evaluated the environmental parameters necessary for handling whole blood for the IFN-γ ELISA and assessed the validity of positive controls. The conclusion was blood should be maintained at room temperature, stimulated within 8 hours and phytohemagglutinin A (PHA) was the most accurate positive control tested in cattle. Finally the last paper attempts to further characterize the meaning of a positive response to the skin test or the IFN-γ ELISA by assessing the consequences of environmental exposure to dead MAP. This study found that inhaled or ingested dead MAP does not stimulate a detectable response on the skin test or IFN-γ ELISA.
CHAPTER 1: INTRODUCTION

1.1 Dissertation organization

This dissertation is composed of an introduction, one literature review chapter, four manuscripts and a general summary. The literature review is a focused review on the history of diagnostic tests based on the cell mediated immune system for paratuberculosis. The first manuscript entitled, “Sensitivity and specificity of the agar-gel-immunodiffusion test, ELISA and the skin test for detection of paratuberculosis in United States Midwest sheep populations” was published in Veterinary Research. The second manuscript entitled, “Evaluation of the gamma interferon ELISA in sheep subclinically infected with Mycobacterium avium subspecies paratuberculosis using a whole cell sonicate or a johnin purified protein derivative” was published in Journal of Veterinary Diagnostic Investigation. The third manuscript, “Time delay, temperature effects and assessment of positive controls on whole blood for the gamma interferon ELISA to detect paratuberculosis” was published in Journal of Veterinary Medicine B, Infectious Diseases and Veterinary Public Health. The fourth manuscript, “Skin test and gamma interferon enzyme-linked immunosorbent assay results in sheep exposed to dead Mycobacterium avium subspecies paratuberculosis organisms” was published in Journal of Veterinary Diagnostic Investigation. The references cited in each chapter are listed at the end of the respective chapters. The general summary chapter discusses the implication of the research presented in this dissertation, weaknesses of the study designs and methods and suggests directions for future research. Appendices are added to support comments made in the general discussion.
1.2 Research summary

Using the cell mediated immune system (CMI) as a diagnostic tool for paratuberculosis is not new, but many researchers have dismissed these tests because they have not correlated with clinical disease. However, serologic tests based on antibody and fecal tests detecting the organism are insensitive so diagnostic tests that can detect animals earlier in the infectious process are needed. Therefore, a new look at CMI based tests using newer more sophisticated study designs and techniques may enhance our understanding and identify potential uses of these diagnostic tests.

Because accurately identifying naturally infected animals is so difficult, the first study uses a Bayesian statistical technique that allows estimates for sensitivity and specificity without the need for a gold standard (accurately classifying animals as infected or non-infected). The Bayesian model used allowed for dependence between the agar gel immunodiffusion test (AGID) and the enzyme-linked immunosorbent assay (ELISA) and considered the skin test, a measure of CMI, conditionally independent. The work was done in collaboration with Ian Gardner at UC Davis. This work showed the skin test was relatively specific (98%) in non-infected populations, and far surpassed the sensitivity of the antibody detection tests, the AGID and ELISA, (72% vs. 8%).

The next study considered a more recently developed CMI diagnostic test, the gamma interferon (IFN-γ) ELISA. Because results of this test are reported as a continuous number, receiver operating characteristic (ROC) analysis was done using bacterial culture of tissues as the reference test. The difficulty in accurately identifying all infected animals using culture was also taken into consideration and a second population of known non-infected animals were also used in this study. The results of this study were similar to the first study. The
gamma interferon ELISA had a sensitivity point estimate of around 67% and a specificity of 98% using sheep from the non-infected population and a specificity of 93% using culture negative animals from the infected population. The difference in specificity between the 2 populations raises questions about the sensitivity of bacterial culture of tissues as the reference test, or residual CMI responses in previously infected or exposed animals. This question is partially addressed in the fourth study.

Whole blood used in the IFN-γ ELISA is sensitive to time delay and temperature while being transported to the laboratory. The third study looks at the effects of various temperatures and time delays on IFN-γ responses in test positive cattle. This study also investigated potential positive controls that could be used in the test to verify if the whole blood has been handled correctly during transport. The results suggested that samples be stimulated within 12 hours after collection and maintained at a temperature between 15-21°C. Unfortunately a true positive control was not identified but phytohemagglutinin A mimicked the response to johnin more closely than any other mitogen or super antigen tested.

The fourth study tested one of the hypothetical causes of a positive CMI test in culture negative animals. Study designs that use bacterial culture of tissues as the reference test for “true test negative” animals instead of animals from non-infected populations often results in lower specificity estimates. One of the potential causes of CMI test positive but culture negative animals is exposure to dead organisms in the environment. In this study sheep from a non-infected population were either exposed to dead organisms orally, via inhalation or left as controls and then tested periodically up to 12 months post exposure. The results suggest that animals do not have CMI responses from exposure to dead organisms and
actual invasion and infection must occur before a CMI response beyond the false positive rate is detected.

1.3 References


CHAPTER 2. REVIEW OF THE LITERATURE

2.1 The discovery of paratuberculosis and development of johnin

The first recognition of Johne’s disease is often credited to Johne and Frothingham in their publication in 1895. However, Søren Saxmose Nielsen of Denmark in his careful review of the early literature makes a compelling case that the disease called “shooters” recognized in by Cartwright in 1829 and by several others, was paratuberculosis.

Nonetheless Johne and Frothingham described the first case that later readers could definitively determine as paratuberculosis. They carefully described the emaciation, thickened intestine and granulomatous inflammation. At the time, Johne and Frothingham thought they had written up an unusual case of tuberculosis, which was endemic in Europe at the time even tough they had failed to demonstrate disease in inoculated guinea pigs from this cow.

In 1906 Bernard Bang determined that the disease was indeed separate from tuberculosis by feeding intestinal scrapings to calves. In 1909 his son O. Bang reported that cattle with paratuberculosis reacted to tuberculin prepared with *Mycobacterium avium*, but not with *Mycobacterium bovis*. It seemed natural to pursue similar diagnostic test development for Johne’s disease that was occurring in tuberculosis (tuberculin testing), as the two diseases were often compared and contrasted in the early literature.

Male in England took note of Bangs’ work and evaluated avian tuberculin intravenously in 2 herds. He saw a significant rise in temperature in some animals that expressed clinical signs of Johne’s disease. Male was the first of many to express frustration in the lack of ability to definitively determine if test positive cattle had Johne’s disease. He
ends his paper with the following. “I am in hopes however, that with the avian tuberculin test the disease will be discovered in its early stages, and then treatment may be commenced with much greater hopes of success.”

Also in 1911 Twort and Ingram reported they were successful at culturing the bacteria from tissues using media prepared for *M. bovis* that included dead *M. phlei* organisms. This critical step allowed the organism to be grown for the production of johnin. At the time there was great hope that a johnin would be more specific than avian tuberculin.

In 1916 M’Fadyean *et. al.* reported they were able to make the first antigen, which he called “johnin” by growing the culture on a liquid medium composed of beef broth using “extracts of other acid fast bacteria.” After the growth of the organism, the organisms were filtered off and the culture media was heat concentrated to form the first johnin. His paper, 83 pages long, describes in detail the use of johnin on cattle, sheep and goats both experimentally and naturally infected from 1911-1916. Each animal is described individually with no summary using descriptive statistics evident in the paper. The objectives of the study were to see whether or not animals would react to johnin and if the reaction was specific. (They were testing specificity only in relation to bovine tuberculosis and avian tuberculosis.) They tested animals intravenously, intradermally, and conjunctively. They found that most of the animals infected with *M. avium* subspecies *paratuberculosis* (MAP) responded to the johnin. However the response to johnin when used on an animal infected with MAP was weaker than the response to tuberculin in animals infected with *Mycobacterium bovis* or avian tuberculin in animals infected with *Mycobacterium avium*. This observation was repeated with a histologic comparison in 1950. M’Fadyean *et. al.* also reported that the
johnin was capable of eliciting a response from animals infected with *M. bovis/M. tuberculosis* or *M. avium*. Thus, their hope that their johnin would be more specific for Johne’s disease than avian tuberculin did not come to fruition.

### 2.2 The use of johnin and avian tuberculin intravenously in the United States to diagnose paratuberculosis

In 1918 Hastings *et al.* at the University of Wisconsin reported the first intravenous use of johnin in the United States. Animals would be injected intravenously with johnin and then their body temperature was monitored over the next 24 hours. Forty-five animals were tested and four animals reacted with an increase in temperature. One clinically-affected animal did not have an increase in temperature. These five animals were examined at post mortem and four were found to be infected; however, no johnin test-negative animals were examined. The authors did not make any conclusions other than they hoped to retest this herd and others around the State. No subsequent follow-up data was ever published.

Two other case reports were published in the United States on the use on johnin IV in cattle. The first was in a herd of 38 cows. This study had similar flaws to previous studies, ie., no test negative cattle were evaluated post mortem and there was a lack of a definitive test for Johne’s disease in animals that were positive on the johnin test and negative on histologic exam. The second case study was a little more thorough. It included a herd of 18 cows, 9 of which were tested twice. Ten of the 18 were examined postmortem both grossly and histologically. Only 2 of the 18 tested negative with johnin and both were negative at necropsy. Sixteen cows tested as suspects or positives with johnin and of these, 8 were examined on necropsy and 3 of these were positive for paratuberculosis.
By the late 1920’s intravenous testing with johnin was the test of choice. There were still concerns about the specificity of johnin and researchers still questioned if johnin was superior to avian tuberculin. Hagan at Cornell compared commercial avian tuberculin to several lots of johnin in a variety of experimentally-infected guinea pigs, naturally-infected tuberculosis chickens and experimentally-infected sheep. He concluded that there was definitely batch to batch variation in johnin lots, and that animals with paratuberculosis were just as likely to react to avian tuberculin as johnin.\textsuperscript{37}

### 2.3 Development of the intradermal skin test for paratuberculosis

The skin test also known as the Mantoux test was first introduced in humans in 1907 for tuberculosis.\textsuperscript{121} It is based on the delayed type hypersensitivity (DTH) reaction that is mediated by cellular immunity. It causes a localized swelling that consists primarily of macrophages, monocytes and T-cells as well as fibrin. This reaction occurs 48-72 hours post injection.

In England, Dunkin gave a detailed report on the manufacture of johnin as well as a description of double intradermal injections, meaning animals would get two injections of johnin in the same area 24-48 hours apart and evaluated for a DTH reaction similar to what was being promoted at the time for diagnosis of tuberculosis. Dunkin was still using a non-synthetic medium and observed non-specific reactions in cattle thought to be uninfected. However he felt these could be differentiated based on size and diffuseness of the swelling.\textsuperscript{26} Isherwood in 1929 followed up Dunkin’s work and reported on the use of the double intradermal skin test, putting significant emphasis on the character (tender diffuse and hot) rather than the size of the swellings.\textsuperscript{44}
In 1935 Minett of England published a study where they tested and followed cattle to slaughter from 7 farms.\textsuperscript{93} These animals were tested with johnin every 3-4 months. The authors were not clear which animals were subject to post mortem testing. A total of 499 animals were tested with 16\% testing positive, 72\% no reaction and 12\% testing suspect. A total of 52 animals were examined at post mortem, 46 of them were reactors suggesting the authors did not examine a representative number of non-reactors. The authors found 6 animals with Johne’s disease that did not respond to the test and 7 animals that responded to the test that were negative on post mortem. One comment the authors made in this longitudinal study is “Repeated tests with johnin on a number of animals showed much variation in the degree of local response, a result which is attributed to variations in the allergic state.” This is the first documentation in the literature of variation on the repeatability of the skin test, a problem that would be mentioned by researchers in several future publications. Also in 1935 McConnell published a case report using johnin in a double intradermal test suggested by Dunkin. While his results are difficult to interpret, his conclusion was the double intradermal test was satisfactory for diagnosing animals in the preclinical stages.\textsuperscript{86}

Johnson and Pratt in 1944 attempted to characterize skin test positive cattle in a herd believed to be non-infected.\textsuperscript{49} They identified a small herd of 35 Holsteins with no history of Johne’s and skin tested them. Seven of the cows were test positive. The cattle were moved to another location and managed as a separate herd. They were tested every 3-4 months for 3 years. Detailed production records on the reactor herd were kept with an annual milk production per animal of 7863.4 lb, which they concluded was desirable production. Throughout the study there were no animals clinically diagnosed with paratuberculosis. A
total of 6 animals were killed or died during the study and 3 were found to have acid-fast bacteria in their intestines with thickened mucosa; however, they were unable to grow the organism. Johnson gave 3 potential reasons why animals reacted to the skin test: 1) active infection was sufficiently severe to bring about characteristic clinical signs and lesions; 2) a less virulent form of the organism may be causing histological lesions but no clinical disease; 3) other mycobacterial infections that do not cause overt disease cause the animal to respond to the skin test.

### 2.4 Improving the growth and production of johnin

Hagan’s work in 1928 showing no difference between johnin or avian tuberculin to diagnose paratuberculosis did not stop the search for a better quality johnin and in 1932 a report was published about making johnin using synthetic media with the addition of dead *M. phlei* organisms.\(^\text{14}\) This report also compared 2 tests for Johne’s disease, intravenous (or subcutaneous) use of johnin combined with temperature monitoring which was more popular at the time than the intradermal test. The authors concluded that both tests were comparable and they were more likely to find evidence of the disease at necropsy if both tests were positive.

Dunkin in 1933 and Watson in 1935 reported growing the organism on synthetic media, thus confirming Broerman’s work in 1932.\(^\text{28,142}\) This was considered a significant advancement because it eliminated potentially reactive animal protein from the media and subsequently from the johnin. Dunkin also reported significant growth differences between strains. He mixed the strains together and grew them on liquid media with and without *M. phlei*. Although he had meager growth without *M. phlei*, he reported the potency of the
subsequent johnin was similar. Others must not have agreed with him as *M. phleci* was a routine ingredient in the manufacturing of johnin until the isolation of mycobactin J.

Throughout the late 30’s and 40’s a significant effort was put forth to improve johnin production and evaluation, particularly with the conversion of heat-concentrated tuberculins to purified protein derivatives (PPD’s) and development of the guinea pig model.32,33,50-53,69,87,93,117,142 Little attention was given to actually characterizing the performance of the skin test. Researchers and practitioners understood that some animals in the late stages of the disease were anergic and showed no reaction to the skin test. They also did not expect that animals would show changes on post mortem exams immediately after skin testing. To diagnose tuberculosis, researchers would grind up lymph glands and inoculate guinea pigs.22,135 Many cases of tuberculosis with no post mortem changes were diagnosed by guinea pig inoculation. Unfortunately a laboratory test animal for paratuberculosis had not been discovered (and has not been even to this day). Consequently, much of early researchers’ and practitioners’ beliefs concerning the efficacy of the skin test for Johne’s disease was based on experience with the use of skin testing for TB rather than actual data demonstrating the efficacy of the skin test for Johne’s disease. E.P. Johnson writes. “The development of a diagnostic agent (johnin) that has a high degree of specificity in detecting these early cases so that the problem may be removed is the goal of the workers on this problem. When achieved, it will place the eradication of Johne’s disease on the same basis as bovine tuberculosis. It is toward that goal that work is being directed by the United States Regional Animal Disease Research Laboratory in Auburn, Alabama.”49

In the 1940’s the Bureau of Animal Industry hired a statistician to assist researchers in designing and evaluating experiments. Wadley published 2 articles on methods of
evaluating and comparing allergens.\textsuperscript{138,139} These experimental designs and methods were used in several studies headed by H. W. Johnson and A. B. Larsen. To reduce measurement variation in these allergen comparison experiments a dermal thickness gauge was introduced and used by the Bureau of Animal Industry in the 1940’s and 50’s.\textsuperscript{73} Also, 2 studies designed to evaluate the effect of intradermal injection locations on the animal were published. These studies found the neck to be much more sensitive than the caudal fold.\textsuperscript{9,76}

Even though McIntosh and Konst reported making purified protein derivatives using MAP in 1943, the Bureau of Animal Industry in the USA continued to use heat- concentrated culture filtrate until 1955 when BAI researchers conducted a study of their own which clearly showed the superiority of PPD’s to heat concentrated culture filtrates.\textsuperscript{75}

Also in the 1950’s researchers began using more sophisticated laboratory methods to fractionate tuberculin and johnin to improve sensitivity and specificity.\textsuperscript{5,7,8,54-56} Researchers used acid, alcohol and phenol to fractionate and purify proteins and then compared the fractions to complete johnin in sensitized cattle. This work showed promising results, however there is no evidence that this information was ever utilized in future experiments and field trials using johnin.

2.5 The consequences of repeated skin testing with johnin

A concern with the skin test for tuberculosis was the anergy created by repeated testing. In 1949, Larsen reported that cattle experimentally sensitized with MAP were desensitized and failed to react to johnin 45 % of the time 1 week post initial testing. He reported that by 4 weeks only 13 % displayed anergy to johnin.\textsuperscript{77} In 1951 Sikes showed that cattle were more susceptible to desensitization if they were naturally infected rather than
artificially infected and desensitization lasted more than 4 weeks. Then in 1953 Sikes followed up this work with a more detailed study where he used eight naturally MAP-infected cows and injected johnin intradermally at 9 different sites per cow and repeated those tests at 14, 16, 18, and 20 weeks post initial test. He found that cattle were more likely to be refractory at 14 weeks post testing (27%) than at 20 weeks (6%). His conclusions were: infected cattle will react to the intradermal johnin test at a previously unused site at 14 weeks, and there was local desensitization when repeated intradermal johnin tests were made in the same site so if the same site is to be used, 20 weeks should pass before the test is used again.

2.6 Comparison of the skin test and the complement fixation test

In 1957 Howell reported he was able to elicit a skin test response in previously negative guinea pigs by passive transfer of cells from the peritoneal cavity of previously sensitized guinea pigs. The transfer of sera from positive guinea pigs did not elicit a skin test response. He concluded that the skin test response to Johne’s disease was due to cellular immunity rather than humoral immunity as was the case for the complement fixation test. Therefore, the disagreement among these two tests in the field was the result of different immune responses. Consequently, Howell felt both tests would be considered valuable when testing for Johne’s disease. This is the first evidence that the skin test response to johnin was mediated by the cell-mediated immune system in the literature.

The interest in using both complement fixation testing and skin testing was relatively high after Scharf (as reported by Howell, 1957) reported that animals positive to both tests were more likely to be found infected at necropsy than animals testing positive to just one
test. To alleviate the concern of the two tests interfering with each other, the Canadian group conducted a study looking at the interference of the skin test to the complement fixation test at 48 hours prior to skin testing and at skin test reading and found skin testing did not interfere with the complement fixation test. In another study in 1964 Pearson observed a rise in titers after skin testing, but he did not have a control group to take into account disease progression.

Several other researchers began evaluating the two tests. Rice et al, 1958, looked at specificity of both tests in orally and intraperitoneally MAP-inoculated cattle. While she did not describe the antigen preparations used in her paper, she found that orally infected animals were less likely to have a reaction to M. avium antigens using the complement fixation test than they were to the johnin test. The cattle inoculated intraperitoneally tested positive to M. avium in both assays.

In Denmark two other groups of researchers attempted to compare the skin test and the complement fixation test in observational field studies. They both had similar study designs in that they parallel tested all cattle in multiple herds then classified the herds according to the complement fixation test results and then observed the distribution of avian tuberculin reactors among the groups. The difference between the two studies is that one study population had high incidence of tuberculous poultry on the farms and the other did not. In the population with high avian tuberculosis, about 15% of the animals tested positive vs. the low incidence population where approximately 3% tested positive. There was no difference in the distribution of avian reactors in the various groups classified according to their complement fixation results in either study. A significant problem with these observational studies was no attempt was made to estimate the prevalence of Johne’s disease
in either of these two populations which would be critical to understanding the performance of the tests.

In 1962 a British Veterinary Medical Association advisory board published a guide to Johne’s disease diagnostic tests. This report looked at test methods and proposed recommendations on diagnostic tests useful for diagnosing the clinical animal and the subclinical animal. They concluded the complement fixation test and examination of acid fast clumps in the manure were the most reliable diagnostic tests for clinical disease. For subclinical disease they dismissed these two tests as not useful. They reported that the johnin skin test held the most promise, but the test was plagued by false positive results. False positives were defined two different ways; animals from infected herds that tested positive with no evidence of disease, and animals testing positive from herds believed to be non-infected. This report considered fecal culture as a potential diagnostic test; however culture methods were not yet that advanced.

2.7 Field utilization of the skin test in tuberculosis vs. paratuberculosis

An important point to consider is that up to and including the 1960’s, tuberculosis and paratuberculosis research were on very similar paths. Two very good reviews of the tuberculin skin test and problems with specificity that discuss the history up to and including the 1960’s are available elsewhere. Just as with johnin, researchers expressed frustration with non-specific reactions and the lack of progress in the laboratory to improve tuberculin beyond the discovery of TCA precipitation for making PPD’s. These tuberculosis researchers in the 1960’s expressed the desire for a serologic test, similar to Johne’s disease researchers. In spite of the similarities in the research programs, applied use of these
diagnostic tests were dramatically different. Skin testing was routine for tuberculosis diagnosis and was readily adopted into official control and eradication programs, with significant success.\textsuperscript{6,30,107,146} But only a few countries adopted skin testing for Johne’s disease in any official capacity and those that did, later dropped programs using skin testing. Only a few countries residually use skin testing for Johne’s disease as an import requirement.\textsuperscript{1}

2.8 Field trials correlating evidence of paratuberculosis with skin test positivity in cattle

E.P. Johnson in 1950 reported limited success in using the skin test to eradicate the disease in a heavily infected Jersey herd.\textsuperscript{48} He followed the herd for over two years and tested it roughly every three months. A total of 46 animals out of 106 were removed and slaughtered. After removing nearly half the cattle and still having a significant number of new reactors at each new test period the authors concluded that cattle were becoming infected at a higher rate than can be justifiably removed. He found no false positive reactors if the induration was at least double the normal skin thickness. His conclusion was management changes would be needed to help decrease the number of new infections in young stock and that the skin test could not be used in heavily infected herds without management changes.

In 1958 Konst and McIntosh reported on the use of johnin PPD in Canada in a series of two articles.\textsuperscript{67,68} They reported on the used of johnin PPD on 8 herds. They suggested they had eradicated Johne’s disease in 4 of the 8 using the skin test. However, no rigorous postmortems were done on cattle not testing positive to Johnin. The authors made their conclusion based on the disappearance of clinical signs for 2-3 years prior to publication.
Correlating skin test results to detectable disease was the overall objective of the next three longitudinal studies. Two of these are perhaps the most important studies conducted in the 1960’s because the results of these studies fundamentally changed the focus of Johne’s disease diagnostic test research in the United States from skin testing to serologically based assays and fecal culture. The third study published in 1980 from Canada confirmed the results of the previous two studies. All three studies are longitudinal. The first was a five-year study involving one herd (161-195 cows) in Alabama. Cows were skin tested periodically and selected tissues were collected at slaughter and smears were made to look for MAP. Ninety-six cattle were examined and of the 41 that were considered positive on intestinal smear, 21 (51% sensitivity) were skin test reactors. Of the 55 that were considered negative on necropsy, 25 were skin test reactors (55% specificity). They also described the persistence of skin test positivity; 4 cattle were intermittent reactors, 12 were routinely positive for 6 months, and two groups of five cattle were routinely positive for 12 and 36 months. There was no difference in the percent positive at necropsy between these groups. It is important to understand that the authors used an induration cutoff of 3mm. This may have affected the sensitivity as the authors give no information on animals that had smaller indurations. Nonetheless, there was little evidence that skin testing would be useful in adult dairy cattle from a heavily infected herd.

The second study was reported in 1968. Merkal et al. compared the complement fixation test, agar gel immuno-diffusion test, the skin test, direct microscopy of fecal material and fecal culture on a herd of Guernseys. They selected 370 of 1500 that were likely to be sent to slaughter in the next 3 years and tested those animals every 6 months. During the next 3 years, 98 cows were positive on fecal culture. The authors failed to state how many
cows were evaluated at necropsy. They compared the other test results to fecal culture. This made fecal culture the gold standard, automatically making fecal culture the most accurate test. None-the-less, the skin test identified only 78% of fecal shedders and identified 50% of cows that were fecal culture negative and negative on necropsy examination.

The final longitudinal study was performed on a 100 cow dairy herd in Quebec as reported by de Lisle.\textsuperscript{24} For the first two years, all animals that were reactors on the skin test (\textgtr=3\text{mm}) and positive on the complement fixation test were culled. These animals were not followed to slaughter. The rest of the herd was depopulated one year later after two more whole-herd skin tests. One-hundred-two animals were examined and 37 of these were positive on histopathologic exam and/or culture of tissues. The skin test only identified 2 of the 37 positive cows (5\% sensitivity) and falsely identified 4 of the 65 cows (94\% specificity). The authors did publish induration responses less than 3 mm. If one considers those data, 16 of the 33 positive cows responded to the skin test (48\% sensitivity), and 21 of 31 negative cows responded to the skin test (32\% specificity). These results are no better than tossing a coin

In the 1980’s two papers were published in India by Paliwal, et al., comparing the skin test to various serological tests in cattle.\textsuperscript{100,101} There was little description of how the skin test was conducted. Researchers found that the skin test identified cattle that were preclinical, but failed to identify clinical animals. The fluorescent antibody test performed just as well as the skin test on preclinical animals, and was also able to identify clinical animals. Kandavel, again in India in a similar study, found fecal smears were more accurate at diagnosing Johne’s disease than the skin test.\textsuperscript{60} All of these papers fail to identify how the test animals were selected and how animals were considered infected.
In 1990 Kormendy in Hungary found similar results with a field trial; 51% sensitivity and 46% specificity using fecal microscopy as the reference test.\textsuperscript{71}

\section*{2.9 Using the skin test in small ruminants to diagnose paratuberculosis}

There was only one publication on the use of the skin test in sheep in the United States prior to the current work reported in this dissertation.\textsuperscript{84} This work was conducted in Bozeman, Montana using the university sheep flock of 130 ewes in the late 1940’s. Researchers first diagnosed Johne’s disease in 1942 in the flock and in 1945, 4 ewes died of clinical Johne’s. They skin tested the flock and found 30 testing positive. They tested the flock several times without culling all test positive animals. Finally in October of 1945 the flock was depopulated and the farm remained free of sheep until November of 1946 when it was repopulated. The new flock was tested 3 times over the course of the next 4 years and a total of 5 sheep tested positive. On necropsy there was no evidence of Johne’s disease in these 5 test-positive sheep. Their conclusion was that a producer could depopulate and repopulate with non-infected sheep 1 year later and successfully eradicate the disease.

The only other time small ruminants were used in the United States to evaluate the skin test was when the BAI decided to substitute goats for cattle when comparing different johnins and tuberculins to reduce costs.\textsuperscript{74} Goats were found comparable to cattle.

Kulshrestha \textit{et al.}, 1948, in India conducted a cross-sectional study comparing the skin test with the complement fixation test and fecal smear examination.\textsuperscript{72} They considered sheep as infected if they showed clinical signs and/or gross lesions at necropsy. The authors found fecal smear examination superior and the skin test routinely failed to identify sheep clinically infected.
2.10 Fecal culture as a diagnostic test for paratuberculosis

The first isolation of MAP from tissues was in 1911\cite{133} and it took another 30 years before a report was published that suggested that cattle could be diagnosed by fecal culture.\cite{94} Several reports were published that improved culture methods throughout the 40’s-90’s including improving isolation,\cite{18,63,64,79} decontamination,\cite{91} identification of mycobactin J\cite{120,122,123}, and media\cite{58,88,89}. Consequently, fecal culture only matured as a diagnostic test after the skin test fell out of favor. For sheep it is important to understand that fecal culture was not considered a good diagnostic test until the early 2000’s with the validation of sophisticated liquid culture techniques and it has yet to be validated in the United States.\cite{143-145}

2.11 The lymphocyte transformation test

Beginning in the 1970’s another diagnostic assay based on the CMI response was investigated called the lymphocyte transformation assay.\cite{3,15,16,23,46} This test was technically complex, requiring the harvesting of lymphocytes, culturing them with radioactive \textsuperscript{125}I and measuring the subsequent radioactivity. There are several research papers that correlate the skin test with the lymphocyte transformation assay in tuberculosis and the two tests have relatively good correlation.\cite{20,62,124}. One would assume there should be a similar correlation with paratuberculosis, although de Lisle, who found the skin test to perform very poorly had much better results with the lymphocyte transformation test.\cite{23,24} One key difference is he used the lymphocyte transformation test on select younger animals. Another study, using either the culture of tissues and/or histology as the reference test (65 of 90 cattle were
positive) found the lymphocyte transformation test to be 56% sensitive and 84% specific compared to fecal culture which was 29% sensitive and 100% specific.\textsuperscript{16}

2.12 The development of the gamma Interferon ELISA

In the early 1990’s a new method of diagnosing tuberculosis was proposed by a group in Australia using an ELISA to quantify the production of gamma interferon (IFN-\(\gamma\)) in white blood cells harvested from whole blood.\textsuperscript{115,116,147,148,150} White blood cells or heparinized whole blood containing the WBC’s were cultured in the presence of tuberculin. White cells that recognize the mycobacterial antigens secrete gamma interferon into the culture media or plasma. The media or plasma can then be evaluated in a sandwich ELISA format specific for the detection of IFN-\(\gamma\).

The cytokine, IFN-\(\gamma\), is a major component of the cell-mediated immune response, secreted primarily by T-lymphocytes.\textsuperscript{31} The antigen used to stimulate the cells during culture was tuberculin PPD. To differentiate cross reactions of avium type mycobacteria, avian PPD was used at the same time. Comparing the strength of the IFN-\(\gamma\) ELISA optical density response between the cells stimulated with tuberculin vs. avian tuberculin would determine the test result, similar to measuring the size of the induration in the comparative cervical skin test.

Similar to the previous history of the skin test, the IFN-\(\gamma\) ELISA was proposed as a method for early diagnosis of paratuberculosis, first in 1992 then in 1996.\textsuperscript{12,125,136} The first two papers were cross sectional comparisons of the absorbed antibody ELISA with the IFN-\(\gamma\) ELISA. These tests were conducted in naturally-infected herds and the authors did not attempt to evaluate specificity. Their estimates of sensitivity were around 70-93% based on
fecal culture and histology. Stabel in 1996 compared five non-infected and seven sub-clinically and 5 clinically-infected animals. She found the IFN-γ ELISA optical density values were much higher in the subclinically infected cows than the control cows.

Several researchers have used the IFN-γ ELISA to characterize the immune response after Johne’s disease vaccination \(19,35,40,66,96,134\) and experimental infection \(34,36,70,108,127,128,130,140\) in cattle, sheep, goats and bison. It is important to note that all of these authors documented very strong IFN-γ responses post vaccination. After experimental infection, animals had a detectable IFN-γ response by 4 months post infection. In studies where animals were sampled every 2-4 weeks for at least a year, all animals had at least one negative IFN-γ ELISA result, and there was a large amount of variation in the strength of response from sample to sample.

At the turn of the century, Biocor obtained approval for the Bovigam IFN-γ ELISA kit in both Australia and the United States as an adjunct test for bovine tuberculosis.\(^{149}\) While approval of the Bovigam kit has yet to be given for use in paratuberculosis, approval for use in tuberculosis allowed researchers to have ready access to the kit, which they were able to modify and begin evaluating under field conditions.

### 2.13 Impact of the IFN-γ ELISA in diagnosing tuberculosis

The approval of the IFN-γ ELISA has offered veterinarians and physicians an alternative diagnostic test for tuberculosis. The literature is controversial and research studies have been mixed in determining what test is the most sensitive.\(^{4,21,65,80,106,111}\) The IFN-γ ELISA has provided an efficient method of testing and evaluating individual and mixtures of proteins as potential antigens. This has led to the characterization of proteins that can
differentiate BCG vaccinated humans from infected humans.\textsuperscript{85} Specifically, ESAT-6\textsuperscript{104,105} and CFP-10\textsuperscript{25} have improved the specificity of the gamma interferon ELISA significantly, both individually and in combination.\textsuperscript{13,137} Later a fusion protein was developed combining the two antigens that induced a stronger response than each protein individually without a significant loss of specificity.\textsuperscript{2,39,141}

Similar research in paratuberculosis is just beginning. Specific proteins are being tested and these may improve the specificity and prevent cross reactions with \textit{M. avium} organisms.\textsuperscript{29,95,98,99,112,114}

\textbf{2.14 Field trials correlating evidence of paratuberculosis with the gamma interferon ELISA in cattle}

Stabel \textit{et al.} were one of the first to publish field data using the Bovigam kit in 2001.\textsuperscript{126} Researchers in this study compared the IFN-\(\gamma\) ELISA to fecal culture (or historical fecal culture results) in 4 dairy herds infected with MAP. These researchers chose to use an OD response cut point similar to what tuberculosis researchers were using, which was 0.1 above the nil sample. Using this criterion, the authors were able to obtain sensitivities ranging from 50-70\% using a single IFN-\(\gamma\) ELISA result. The study design used did not allow for an evaluation of specificity.

Paolicchi \textit{et al.} in Argentina evaluated the IFN-\(\gamma\) ELISA, the antibody ELISA, milk and fecal culture and the AGID by randomly selecting 24 cows from an infected dairy and then running the tests one time.\textsuperscript{102} They used culture results as the gold standard. A total of 9 animals were culture positive, 7 from feces and 2 from milk. The IFN-\(\gamma\) ELISA identified only 2 of the culture positive cows and the serological ELISA correctly identified 6 of the
culture positive cows. They concluded that the IFN-γ ELISA was not useful in diagnosing paratuberculosis. In 2002 Jungsersen et al. evaluated the IFN-γ ELISA in 10 infected and 5 non-infected dairy herds in Denmark. They evaluated 259 animals from the infected herds and 119 animals from the non-infected herds. They did not calculate sensitivity, but estimated the specificity of the IFN-γ ELISA to be 95-99%. In particular they found more false positives in calves less than 15 months of age.

Kalis et al. investigated the specificity of the IFN-γ ELISA and the skin test for Johne’s disease in the Netherlands. They used 35 dairy herds believed to be non-infected selected from the Dutch paratuberculosis-free certification program. They sampled 20 younger animals (6-24 mo) and 20 animals > 24 mo on each farm. They found significant differences between lots of PPD on both assays (96%, 81%, 99.5% specificity for the IFN-γ ELISA and 92% vs. 95% for the skin test). If they used the results from all antigens, the skin test was 88% specific using an induration cut point of 2mm, 91% using a cut point of 3mm, and 93.5% using a cut point of 4mm. The IFN-γ ELISA was not specific (66%) using the cut point of .1 above the non-stimulated control. Kalis developed a new algorithm and using that, they were able to achieve a specificity of 93.6%. The authors were also able to document significant herd to herd variation in specificity that they were able to associate with different kinds of bedding. The specificity for the skin test ranged from 58% to 100%.

In 2003 Huda et al. published two papers on the performance of the IFN-γ ELISA in dairy cattle. In the first paper they followed 16 dairy cattle from infected herds to slaughter. Six were positive on necropsy. These cattle were subjected to serial testing 3 times approximately 2-6 months apart. All six of the infected animals had at least two positive IFN-γ responses, but three had one negative response. Three of the 10 non-infected cattle had one
positive response to the IFN-γ ELISA. In the second paper the authors sampled 363 dairy cows (247 from infected herds and 116 from non-infected herds) three times 5-8 weeks apart. Four percent of the samples were considered invalid due to high background levels of IFN-γ and other reasons. Three percent of the animals in the non-infected herds tested positive on the IFN-γ ELISA at least once, a specificity of 97%. Forty-eight percent tested positive at least once from the infected herds, but only 21% tested positive at all 3 samplings and 14% only tested positive one of the three times. These results document that the IFN-γ ELISA lacks consistency or repeatability, similar to the skin test.

In 2004 Huda et al. reported on a longitudinal study using the same herds and animals that were in her previous articles, but over a two year period.43 The authors divided the cattle into infected (based on culture), exposed (culture negative animals in infected herds) and non-infected cows (animals from non-infected herds) and used Receiver-operating analysis (ROC curves) to evaluate the accuracy of the IFN-γ ELISA as well as the milk and serum antibody ELISA. They found that the IFN-γ ELISA was the most accurate test when cattle were less than 3 years of age. After 3 years of age, there were no significant differences between the diagnostic tests. The specificity was similar between the tests if the gold standard used for non-infected animals was cows from non-infected herds rather than culture.

**2.15 Using the IFN-γ ELISA in small ruminants to diagnose paratuberculosis**

In 1999 Burrells was the first to report the use of the IFN-γ ELISA in sheep.17 He used the test to demonstrate a relationship between the cell-mediated immune response and histological lesions. He found that sheep having multibacillary lesions in the intestines had lower gamma interferon responses than sheep with paucibacillary lesions.
In 2003 Manning et al. used the IFN-γ ELISA in pigmy goats.\textsuperscript{83} Thirteen goats from a herd of 30 were selected after testing positive on either fecal culture or the antibody ELISA. These goats were then serially tested with the IFN-γ ELISA, fecal culture, and antibody ELISA. They found the IFN-γ ELISA identified 7 of the 10 goats that were found to be infected at necropsy and 1 of the 3 goats that were non-infected based on necropsy results.

In 2006 Storset evaluated the IFN-γ ELISA in vaccinated goat herds with paratuberculosis and non-vaccinated goat herds without paratuberculosis.\textsuperscript{129} While vaccination confounded the results in infected animals, only 3 of the 376 animals tested positive in non-infected herds. This included 255 goats less than 18 months of age suggesting that non-specific reactions associated with age that have been reported in cattle were not a problem in goats.

### 2.16 Research objectives

Although the IFN-γ ELISA is a USDA-approved kit, no standard methods of interpreting the results for paratuberculosis have been suggested by the manufacturer and several different methods have been proposed in the literature, with no consensus, particularly in sheep. Also of concern is the fragile nature of whole blood used in the assay and lack of published data on transportation requirements and positive controls.

Furthermore, novel ways are needed to evaluate the sensitivity and specificity of the skin test and the IFN-γ ELISA that lessen the problem of a reference test or gold standard to measure the performance of these diagnostic tests. The objectives of the studies described in this dissertation were to: 1. Estimate the sensitivity and specificity of the skin test, antibody ELISA, and agar gel immuno-diffusion test in two sheep populations with out a gold
standard. 2. Estimate the sensitivity and specificity of the IFN-γ ELISA using ROC analysis.
3. Characterize the effects of temperature and time delay on whole blood used for the IFN-γ ELSISA. 4. Evaluate the effects of exposure to dead organisms on the response in the skin test or IFN-γ ELISA.

2.17 References

11 Bang O: 1909, Das geflügel tuberkulin als diagnostische mittel bei der chronischen pseudotuberculosen darmentzundung des rindes (Johne's disease)[Avian tuberculin as a diagnostic test for the chronic pseudotuberculous enteric inflamation of cattle]. Centbl Bakt Parasit Infkr Abt I Orig 51:450-455.


71 Kormendy B: 1990, Paratuberculosis in a cattle herd: Comparison of allergic, serologic and faecal microscopic tests. ACTA Microbiologica Academiae Hungaricae 37:219-222.


Twort FW, Ingram GLY: 1912, A method for isolating and cultivating the Mycobacterium enteritidis pseudotuberculosis bovis, Johne and some experiments on


CHAPTER 3. SENSITIVITY AND SPECIFICITY OF THE AGAR-GEL-IMMUNODIFFUSION TEST, ELISA AND THE SKIN TEST FOR DETECTION OF PARATUBERCULOSIS IN UNITED STATES MIDWEST SHEEP POPULATIONS

A paper published in Veterinary Research*

Suelee Robbe-Austerman, DVM, MS\textsuperscript{a}, Ian A. Gardner, BVSc, MPVM, PhD\textsuperscript{b}, Bruce V. Thomsen, DVM, PhD\textsuperscript{c}, Daniel G. Morrical, PhD\textsuperscript{d}, Barbara M. Martin, MS\textsuperscript{c}, Mitchell V. Palmer, DVM, PhD\textsuperscript{a}, Charles O. Thoen, DVM, PhD\textsuperscript{e}, Chad Ewing, DVM\textsuperscript{f}

*Reprinted with permission of Veterinary Research, 2006, 37:553-564

\textsuperscript{a} National Animal Disease Center; Agricultural Research Service, USDA; 2300 Dayton, Ames, Iowa 50010, U.S.A

\textsuperscript{b} Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, CA 95616, U.S.A

\textsuperscript{c} National Veterinary Services Laboratories, USDA; 1800 Dayton; Ames, Iowa 50010 U.S.A

\textsuperscript{d} Department of Animal Sciences; Iowa State University; Ames, Iowa 50011 U.S.A

\textsuperscript{e} Veterinary Microbiology and Preventive Medicine; Iowa State University; Ames, IA 50011 U.S.A

\textsuperscript{f} Groton Veterinary Clinic; 1503 N Broadway; Groton, SD 57445 U.S.A

Funding sources: Iowa Healthy Livestock Advisory Council, USDA/ARS CRIS Project 310-062, and the Johne's Disease Integrated Program in Research, Education, and Extension (USDA-CSREES award #2003-05165)
3.1 Abstract

Our objective was to estimate the sensitivity and specificity of the agar-gel-immunodiffusion test (AGID), the antibody ELISA, and the skin test for the detection of *Mycobacterium avium* subspecies *paratuberculosis* (*MAP*) in sheep using Bayesian methods without a gold standard. Fourteen flocks (2465 sheep) were used. Five flocks (450 sheep) were considered *MAP* non-infected and 9 flocks (2015 sheep) had sheep infected with *MAP*. Sheep were skin tested and blood was collected for AGID and ELISA testing. Results were analyzed using a Bayesian 3-test in 1-population model fitted in WinBUGS. The model allowed for dependence (correlation) between the two serologic tests, but these two tests were assumed to be conditionally independent of the skin test. The estimated specificity was 99.5% (95% PI of 98.9-99.9%) for the AGID; 99.3% (98.4-99.8%) for the ELISA using an optical density measured cutoff of 0.20; 99.2% (98.1-99.8%) using a cutoff of 0.15; 97.5% (95.8-98.7%) using a cutoff of 0.10; and 98.7% (97.3-99.5%) for the skin test. The estimated sensitivities were 8.3% (6.2-10.7%) for the AGID; 8.0% (6.0-10.4%) 10.6% (8.3-13.1%), and 16.3% (13.5-19.4%) for the ELISA using the cutoffs 0.20, 0.15, and 0.10 respectively; and 73.3% (62.3-85.8%) for the skin test. The skin test was specific in non-infected populations and sensitive in infected populations, although in some cases a positive skin test might represent *MAP* exposure rather than infection. The AGID and ELISA were specific but lacked sensitivity. The AGID and ELISA consistently identified two different populations of infected sheep with only moderate overlap between positive test results.
3.2 Introduction

Mycobacterium avium subspecies paratuberculosis (MAP) causes Johne’s disease which is characterized by weight loss and premature culling/death in all ruminant species. Incubation period is typically years with not all animals developing clinical signs.

The two most common serologic tests available are the agar gel immunodiffusion (AGID) test and the absorbed ELISA. One ELISA (Parachek™, Biocor Animal Health, Omaha, NE, U.S.A.) is approved for use in sheep in the U.S.A. Several studies have evaluated the sensitivity and specificity of the AGID and ELISA in sheep.[5, 14, 28] The AGID usually has slightly higher specificity (99-100%) than the ELISA (95-100%) but this depends on the gold standard used and the ELISA cutoff chosen. The sensitivity of both tests is dependent on stage of infection, with the highest sensitivity in sheep with multibacillary lesions and poor body condition and lowest in sheep in the early stages of infection.[25, 28] Overall sensitivity has been estimated to be low for both tests, from 14-62% depending on the gold standard used and the population tested. The majority of reports suggest a sensitivity between 20-30% if no consideration is given to body condition.[5, 14, 24, 28, 30]

There are no published studies that have evaluated the ELISA in U.S.A. sheep populations. Most of the studies evaluating diagnostic tests in sheep cited above used populations consisting mainly of one breed and similar production systems, which is reflective of the sheep in those environments. In the U.S.A., the sheep industry is highly diverse with many different and often unusual breeds, as well as very different production systems.

Because antibody tests lack sensitivity, especially in early preclinical stages of infection, and the cell-mediated immune (CMI) response is thought to dominate in such
animals, investigations of CMI-based diagnostic tests are warranted. The skin test and the interferon gamma (IFN-γ) ELISA are two such tests that may have application to the diagnosis of MAP infection. Several investigators have used the skin test [12, 16, 19-21, 26, 27, 36, 38] and the IFN-γ ELISA [1, 3, 13, 26, 27, 32] as research tools to characterize immune response and some have promoted these tests as potential diagnostic tools especially in the early phases of infection.

Validation of CMI-based diagnostic tests for paratuberculosis infection can be particularly difficult because a highly accurate reference test for animals in the early stages of infection is lacking. The sensitivity of bacteriologic culture of tissues and/or histology, the reference test which CMI tests have been evaluated against in the past, is affected by sample location and number of tissues taken. [6, 22] Some authors suggest up to 100 different tissue samples may be required to evaluate true infection status. [39] This could partially explain why previous investigators have concluded the skin test lacks specificity when animals in infected populations have been used for specificity determination. [6, 22]

Alternative methods for evaluation of diagnostic tests must be employed when a reasonable reference test or gold standard is not readily available. Recent developments [15] have been made to techniques first suggested in the early 1980’s by Hui and Walter that allow for the evaluation of more than two tests without assuming conditional independence. Estimation of the accuracy of conditionally dependent tests also can be done using Bayesian modeling, which is reviewed in detail elsewhere. [2]

The objective of this study was to estimate sensitivity and specificity of the AGID, ELISA and skin test using Bayesian methods in the absence of a gold standard.
3.3 Materials and methods:

3.3.1 Non-infected sheep population

Flocks owned by Universities or veterinarians located in the Midwest U.S.A. were contacted by the senior author (SRA) as they were likely to have detailed information on health events. Flocks were selected for consideration as potentially free of MAP infection if they had no history of Johne’s disease, conducted routine necropsies on dead and debilitated animals, had not recently purchased females (5 years or longer) and had no purchased females in the flock, and followed management practices (i.e. no contact with cattle, purchased colostrum, etc…) that minimized the risk of introduction or transmission of MAP, if present. Owners of these flocks also agreed to submit all sheep testing positive on AGID, ELISA, or skin test in the study (or a subset of at least 5 if more than 5 tested positive) for necropsy and further testing as described in section 2.6.

Five flocks (450 total sheep) were classified as potentially MAP non-infected based on the described criteria. They were located in Iowa (3), Nebraska (1) and South Dakota (1). Two, (1N, 2N, see Table I) were small flocks with 23 and 28 animals, respectively. These animals were not regularly bred and all offspring produced stayed on the farm; one flock consisted of Jacob and Icelandic sheep, the other Finnish Landrace/Dorset cross. The other 3 flocks were commercial flocks, one (3N) a traditional Midwest winter lambing flock consisting of 119 Columbia/Hampshire cross ewes, another (4N) an intensively-managed accelerated (rams exposed to the ewes every 8 months) lambing flock of 174 Polypay/Dorset/Romanov cross ewes and lastly (5N) a 106 Dorset/Polypay flock that pasture lambed in May. Two of these flocks (4N, 5N) had high prevalences of caseous lymphadenitis (CLA).
3.32 Infected sheep population

*MAP*-infected flocks were identified by diagnostic pathologists from Iowa, South Dakota, and North Dakota based on animal or tissue samples submitted to state diagnostic laboratories. Pathologists contacted owners of these flocks and requested their participation in the study. For those producers who agreed to participate and before they were enrolled in the study, archived formalin-fixed tissues submitted previously to the diagnostic lab were confirmed infected with *MAP* by using a polymerase chain reaction (PCR) test with primers for the IS900 sequence on the formalin-fixed tissues.[23] Flocks that had at least one *MAP* PCR-confirmed sheep fit the criteria to be included in the study as an infected flock.

Detailed histories including production records, culling rate and death loss were collected from each flock to estimate the annual incidence of clinical disease. This information was then used to estimate the likely prevalence of infection for the Bayesian model, described in section 2.7.

Nine sheep flocks, ranging in size from 27 to 502 animals, were included in the study. The annual incidence of clinical disease ranged from 0 to 11% (of total flock size). Two flocks consisting of 173 and 27 sheep were Suffolk cross club lamb (sheep raised for exhibition) flocks (1P, 3P). One was a 502 sheep whiteface crossbred (Rambouillet, Dorset, Finn, Targhee, Romanov) accelerated lambing flock (2P). Three were seedstock, two Suffolk, one with 130 sheep, the other 150 sheep (7P, 9P), and one a 50 ewe Romanov Flock (6P). Three were Rambouillet cross range flocks consisting of 381, 342, and 260 sheep (4P, 5P, 8P). Flocks were located in Kansas (2), Iowa (4), South Dakota (2), and North Dakota (1).
3.33 Skin testing and serum collection procedure

All sheep aged ≥12 months were tested. On day one, animals were injected intradermally with 0.10 ml of Johnin purified protein derivative (PPD) (lot 9801, NVSL/VS/APHIS/USDA, Ames, IA, U.S.A.) in the woolless area of the axillary region. This PPD was produced and evaluated in guinea pigs as previously described.[34] Ten ml of blood were also collected for AGID and ELISA testing. After 72 ± 4 hours, the skin test site was palpated and measured for induration. Because there was minimal variation in skin thickness (mean ± SD = 4mm ± 0.5) in the axillary region from sheep to sheep, pre-measurements were not taken. If there was a palpable swelling at the injection site when reading the skin test, the animal was considered positive and a post injection measurement was taken. Four mm were subtracted from the overall measurement to get the final induration measurement. All animals with palpable inductions were considered positive.

3.34 AGID testing

AGID testing was performed as previously described.[30] Briefly, agar plates were prepared with 0.7% agar (Agarose 1, biotechnology grade, Amresco, Solon, OH, U.S.A.) dissolved in a borate buffer solution (0.2% NaOH, 0.9% H3BO3, pH 8.6) containing 7.0% NaCl. Thirteen ml of agar were placed in 100 mm diameter petri dishes and allowed to cool. Five mm wells were punched in the cooled agar with one center well and 6 surrounding wells placed 3mm from the center well. Paratuberculosis protoplasmic antigen (Allied Monitor, Fayette, MO, U.S.A.) was obtained and diluted adding 2 ml of sterile saline to 10mg of antigen. Forty µl were pipette into the center well. A positive control serum was placed in every other well alternating with the test sera. Plates were read at 24 hours and 48 hours. Tests were
considered positive if a line of precipitation was fully formed between the test well and the antigen well and was continuous with the line formed by the positive control wells.

3.35 ELISA Testing

The ELISA (Parachek™, Biocor Animal Health, Omaha, NE, U.S.A.) was used according to manufacturer’s directions. Briefly, 25 µl of test sera were added to 475 µl of serum diluent buffer for a 30 minute absorption incubation at room temperature, then 100µl of this mixture were added to the coated microtiter plate and incubated another 30 minutes. Plates were washed 6 times with supplied wash buffer. One hundred µl of diluted conjugant were added to the plate and incubated for 30 minutes, then the washing step was repeated. One hundred µl of freshly prepared enzyme substrate solution were added and plates were read with a 650 nm filter. Stopping solution was added when the OD value of the positive controls read between 0.35-0.40. The final OD values were read at 450 nm. The test was evaluated against three different cutoffs, the mean of the negative controls plus 0.1(the USDA approved cattle cutoff), 0.15, and 0.2 (the USDA approved sheep cutoff).

3.36 Necropsy

Seven sheep from non-infected flocks and 32 sheep from infected flocks (not including the initial necropsy submitted to diagnostic laboratories that allowed flocks to be considered for inclusion in the study) were euthanized with intravenous sodium pentobarbital, 90 mg/kg. Animals were examined for gross lesions and the following tissues were taken for histologic examination: ileo-cecal valve, and associated mesenteric lymph node, distal ileum, proximal ileum, distal jejunum, and associated mesenteric lymph node,
mid jejunum, and associated mesenteric lymph node, proximal jejunum, and associated lymph node, duodenum, and hepatic lymph node. Tissues were routinely processed to paraffin blocks. Sections of 5 µm thickness were stained with hematoxylin and eosin and examined by light microscopy. Adjacent sections were stained by the Ziehl-Neelsen technique to visualize acid-fast bacteria. Animals were considered infected with MAP when a granulomatous enteritis and or granulomatous lymphadenitis with acid-fast bacilli were present. Pathologists were blinded to the antemortem test status of the sheep.

3.37 Statistical analysis

Sensitivity and specificity of the AGID, ELISA, and skin tests were estimated by Bayesian methods using a 3-test 1-population model. The model allowed for dependence (correlation) between the two serologic tests but these two tests were assumed to be conditionally independent of the third test (skin test) because they measured different biological responses. First, data from the five non-infected flocks were pooled into one superflock of 450 sheep. Specificity values were estimated directly from these data and beta (α,β) distributions for the specificity of each test were derived[37] for use in the 3-test 1-population model. Beta distributions provide a flexible means for modeling binomial probabilities in a Bayesian analysis because they are constrained between 0 and 1. The shape of the beta distribution is determined by the relative magnitudes of the values for α and β. The mean and variance of the beta distribution are $\alpha / (\alpha + \beta)$ and $\alpha\beta / ((\alpha + \beta)^2(\alpha + \beta + 1))$, respectively. Because the estimation problem is non-identifiable without additional information, one of the authors (SRA) provided expert opinion about the likely sensitivity values for the 3 tests, following recommended guidelines.[35] The expert-elicited most
likely (modal) value and corresponding beta distributions for each parameter are shown in Table II. Second, data for the 9 infected flocks were pooled into a single flock of 2105 sheep. The prevalence of MAP infection in this hypothetical flock was uncertain but the senior author (SRA) considered the most likely value to be 30% and she was 95% sure that prevalence was <70%. This information equated to a beta (2.13, 3.64) distribution for prevalence.

Dependence between the AGID and ELISA was modeled with a parameterization[7] that specified uniform prior distributions for pair-wise sensitivity and specificity covariances, which quantify the magnitude of dependence between the tests. The covariances have upper and lower limits that are determined by the numeric values of the sensitivities and specificities of the tests.[9] The model was fitted in WinBUGS (MRC Biostatistics Unit, Cambridge, UK) using Gibbs sampling and code adapted from elsewhere.[2,31] Posterior inferences were based on 50,000 iterations after discarding an initial “burn-in” of 5,000 iterations. Outputs from the model were the median estimates and 95% probability intervals (PI) for the sensitivity and specificity of each test, prevalence, and the covariances between the AGID and ELISA. Model convergence was checked by running multiple chains from different starting values.[10]

Sensitivity analyses were done using non-informative beta (1,1) priors for prevalence and sensitivity of the skin test. To examine the effects of change in ELISA cutoff values, the frequencies of various combinations of test results were recalculated at cutoff values of mean + 0.10 and mean + 0.15 instead of the default value of mean + 0.20 for the ELISA and the model was rerun for these new values.
3. 4 Results

3.41 Test results

The test prevalence in each flock infected with MAP correlated with the level of clinical disease observed (Table I). Flock 1P had the highest annual incidence of clinical Johne’s disease (estimated 11%) and flock 9P had the lowest incidence (0%) of clinical disease. Test prevalence increased in all three tests from 1-year-old sheep to 2-year-old sheep. After sheep reached 2 years, test prevalence did not increase (Table III). There seemed to be a slight decrease in sheep older than 4 years of age, however high prevalence flocks had fewer older sheep than the low prevalence flocks and this likely caused the apparent decrease as this pattern did not hold true when test prevalence was compared against age within flock (data not shown).

3.42 Necropsy

All animals in the non-infected flocks testing positive on the ELISA, the AGID, or the skin test for paratuberculosis (7 sheep total, Table I) were necropsied and no lesions suggestive of MAP infection were found either grossly or histologically. Of the 32 sheep necropsied in infected flocks, 10 were clinical animals euthanized while at the farm collecting samples for the study. All but 2 had lesions consistent with paratuberculosis, one had severe disseminated caseous lymphadenitis abscesses, and the other had histologic lung lesions consistent with ovine progressive pneumonia. All but one infected flock had animals with clinical Johne’s disease. In flock 9P, (the only infected flock without a history of clinical Johne’s disease) 22 cull sheep were necropsied over the next 3 years after the study, 3 were skin test positive and 19 skin test negative. Two of the 3 skin test positive sheep were
confirmed to have paratuberculosis on necropsy (tissue PCR) and all 19 skin test negative and one skin test positive sheep did not have lesions either grossly or histologically suggestive of paratuberculosis.

3.43 Sensitivity and specificity

The Bayesian analysis indicated that the sensitivity of the skin test (median = 0.73) was substantially greater than that of either of the two serologic tests (median for both tests = 0.08) (Table IV). The specificities of all three tests were high (≥ 0.98) unless the ELISA cutoff was lowered to the approved cattle cutoff of mean + 0.10. The ELISA sensitivity improved (0.08 to 0.16) by lowering the cutoff, but specificity was also affected (0.99 to 0.975). There was evidence of a positive dependence in the sensitivities and specificities of the AGID and ELISA (data not shown). The posterior 95% intervals for sensitivity estimates were much narrower than the prior 95% intervals. A sensitivity analysis using non-informative priors for prevalence and sensitivity of the skin test did not change test accuracy estimates markedly (data not shown).

3.44 Agreement between AGID and ELISA

Table V compares AGID and ELISA results using different cutoffs for the ELISA from all sheep in the flocks with Johne’s disease. Decreasing the cutoff did not improve agreement between the AGID and ELISA. Only five more sheep were identified as both AGID and ELISA positive when the ELISA cutoff was lowered from 0.20 to 0.10, however the number of sheep positive on the ELISA, but not the AGID also increased from 49 to 146, an increase of 97 sheep.
3.45 Induration size of skin test responses

The skin test induration response is shown in Figure 1. Any palpable response was considered positive and measured. From previous experiments, (data not shown) indurations of \( \geq 1 \text{mm} \) were routinely palpable. Forty-one percent \((235/572)\) of animals that were skin test positive in infected populations had induration sizes \( \leq 3 \text{mm} \). These small indurations were not detected in the non-infected flocks; the false-positive induration responses in non-infected flocks ranged from 3 to 9 mm.

3.5 Discussion

The diversity of breeds and production systems represented in this study are reflective of U.S. sheep industry. The location however was restricted to the Midwest, and sheep flocks in other areas of the U.S. should be evaluated in future studies. For test evaluation, we used a Bayesian model that assumed that none of the tests was a gold standard. The model allowed for dependence between the 2 serologic tests but assumed that both tests were independent of the skin test. We had precise information on specificity of the assays using test data from the non-infected population and this allowed us to model specificity with a highly informative prior distribution. This indirectly allowed estimation of sensitivity of the 3 tests with reasonable precision.

A problem encountered when evaluating tests in general is the size of the populations needed.\([8, 11]\) The probability intervals around the point estimates can be quite large, considering the number of animals tested and the weight given to the prior information. This is not such a problem if the results are near 0 or 1, such as the specificity estimates, or even
the AGID and ELISA sensitivity estimates in this study; however, the probability interval around the point estimate of 0.73 for skin test sensitivity was moderately large (0.623-0.858).

The skin test in sheep was surprisingly specific (>98%) even though no skin thickness/induration cutoff was used in the present study. This finding is contrary to previous results in cattle, which vary as to what induration size must be present in order to consider an animal positive. Kalis et al. (2003) used an induration of ≥ 4 mm. If they used an induration size of ≥ 2 mm the specificity of the skin test decreased from 93 to 88%. [18] In their study, no mention was made as to what amount of induration was routinely palpable; they measured skin thickness 72 hours post injection at the injection site and 10 cm behind the injection site. Other cattle studies evaluating skin testing used 3 mm as a cut-off between positive and negative using the difference in pre- and post-testing measurements.[6, 22]

Our data clearly show that many sheep in flocks with MAP infection tested positive at < 3 mm of induration. Test results of several of these animals were correspondingly positive with the AGID or ELISA test, making one assume that we would likely miss some infected sheep if we required an induration of ≥ 3 mm. In our study, animals in non-infected populations did not have small palpable indurations; consequently, we felt that any induration in a sheep in an infected flock is most likely to indicate MAP exposure or infection. It is important to note that sheep have much thinner skin than cattle, and one skin testing technique may not be ideal in both species.

Another, possibly more likely, reason for differences in induration in the non-infected populations between other studies and this one could be the antigen used. In our hands, the 9801 Johnin PPD has been more specific than any other PPD we have tried. Specificities have varied over 10% between lots, even with lots made from the same seed culture (data not
shown). Our experience is consistent with previous studies.[17, 18] When and if an antigen for skin testing becomes standardized, it may be possible to assign likelihood ratios and recommendations based on induration size, so a measurement may always be desirable. However, caution must be exercised if one attempts to use a cutoff value recommended in another paper if the antigen is not the same.

The biological importance of a positive skin test result in MAP infected flocks is unknown. In flock 1P, 60% of the sheep tested had positive skin test results. No management changes were made and the flock continued to have a 7 to 10% annual loss attributable to clinical Johne’s disease in the 3 years post testing, which was the similar to losses in the 2 years prior to testing. The future challenge will be to determine the utility of skin testing and other CMI diagnostic tests for identifying infected individuals in populations where MAP had been diagnosed. Obviously, culling 60% of a flock will have limited acceptance, and even in those flocks where culling is possible, the robustness and repeatability of the skin test reaction must be understood; especially knowing the probability of a subclinically infected animal testing negative.

Because the ELISA and AGID tests lack sensitivity, they usually don’t significantly change negative predicative value, in other words, a negative test for a sheep in an infected flock provides little reassurance to a producer that the animal is truly negative. Consequently, these tests have been promoted as herd tests, more effective at identifying infected flocks rather than individual animals.[4, 29] In this study the AGID and ELISA were both able to identify 8 of 9 infected flocks, even with an estimated (individual animal) sensitivity of only 8%. However, infected flocks were not randomly selected without regard to previous history, so these flocks may have a higher prevalence of paratuberculosis than routinely found. With
the concern of the biological relevance of a positive skin test on an individual animal basis, there is also potential for the skin test to be used as a herd test. With a much higher sensitivity, fewer animals would have to be tested. A major concern with herd level testing is specificity. Even with a specificity of 98.7%, the skin test would have a tendency to falsely identify non-infected flocks, as demonstrated by identifying 3 of the 5 non-infected flocks in this study. A small improvement in specificity would dramatically improve the potential for the skin test to be a useful herd-level test.

In the present study, the AGID and the ELISA had lower sensitivity (both at 8%) than reported in other published studies. There are several possible explanations for this finding. First, approximately 20% of animals tested in our study were between 12 and 18 months of age and many of the other studies evaluated an aged ewe population excluding ewe lambs and hoggets. Second, sensitivity in our study was evaluated using a statistical technique that was not constrained by an imperfect gold standard. Third, it is possible that the latent class for the skin test estimated in the Bayesian analysis might be a mixture of both infection and exposure whereas the AGID and ELISA latent class might only consist of infection. Since all of three tests measure immune response, it is not known if an immune response above the threshold always signifies infection, especially for the skin test and even for the ELISA at lower cutoffs. In herds or flocks with MAP infected animals, there is an opportunity for antigen exposure without infection but the modeling approach cannot distinguish between these 2 states. Because of the model structure that we used, the estimated sensitivity value of 0.73 for the skin test should be interpreted as a weighted average of the sensitivities across the 9 infected flocks rather than a flock-specific estimate.
In the United States, it is difficult to identify non-infected sheep populations, and our population of 5 flocks of 450 ewes was barely adequate to evaluate specificity across a spectrum of management practices and potential exposure to antigenically-related organisms. Our data suggest that ELISA specificity was not significantly affected if the cutoff value for sheep was decreased from 0.20 down to 0.15. This finding should be further evaluated on larger numbers of sheep before the standard cut-off is altered.

It is interesting that the ELISA and AGID identified different populations of infected animals. This was a consistent finding in the larger infected flocks which had multiple ELISA and AGID positive sheep. Several of these sheep that had discordant results were examined at necropsy and found to be infected. This finding has been reported previously.[14, 28] In this study, the agreement between these two tests was not improved by lowering the ELISA cutoff. This only slightly increased the number of AGID positive animals it identified, and dramatically increased the number of ELISA positive animals.

This study highlights the potential for skin testing to detect MAP infection at both the individual and herd level. However, serious obstacles need to addressed, mainly the need to understand the biological relevance of a positive skin test for proper test interpretation and the ability to consistently produce a quality antigen.

3.6 Acknowledgments

The authors would like to thank Dr. Janice Miller for performing the PCR testing.
3.7 References


[34] Steadham, E.M., Martin, B.M., Thoen, C.O., Production of a Mycobacterium avium ssp. paratuberculosis purified protein derivative (PPD) and evaluation of potency in guinea pigs, Biologicals 30 (2002) 93-95.


Table I. The number of sheep with positive and negative test results for various combinations of skin test, AGID and ELISA in 9 flocks with *M. avium* subsp. *paratuberculosis* (*MAP*) infected sheep and 5 non-infected flocks as well as overall test prevalence for the AGID, ELISA, and skin test. The cutoff used for the ELISA was the mean value of negative controls + 0.20.

<table>
<thead>
<tr>
<th>No. sheep</th>
<th>MAP infected flocks</th>
<th>MAP non-infected flocks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1P</td>
<td>2P</td>
</tr>
<tr>
<td>S+, A+, E+</td>
<td>173</td>
<td>502</td>
</tr>
<tr>
<td>S+, A+, E-</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>S+, A-, E+</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>S+, A-, E-</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>S-, A+, E+</td>
<td>81</td>
<td>183</td>
</tr>
<tr>
<td>S-, A+, E-</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>S-, A-, E+</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>S-, A-, E-</td>
<td>7</td>
<td>2</td>
</tr>
</tbody>
</table>

Test prevalence %

<table>
<thead>
<tr>
<th></th>
<th>AGID</th>
<th>ELISA</th>
<th>Skin Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7.5</td>
<td>14.5</td>
<td>60.1</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>4.2</td>
<td>41.8</td>
</tr>
<tr>
<td></td>
<td>3.7</td>
<td>3.7</td>
<td>44.4</td>
</tr>
<tr>
<td></td>
<td>2.6</td>
<td>1.6</td>
<td>28.3</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>1.8</td>
<td>27.1</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>6.0</td>
<td>16.0</td>
</tr>
<tr>
<td></td>
<td>2.3</td>
<td>2.3</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>1.9</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>0.0</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>1.2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>2.5</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>0.0</td>
<td>0.9</td>
</tr>
</tbody>
</table>

\(^a\) S=Skin test, A=AGID, E=ELISA
Table III. Expert-elicited values and corresponding beta (α, β) distributions for sensitivity of 3 diagnostic tests for ovine Johne’s disease

<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity (modal value)</th>
<th>Sensitivity (upper or lower limit)</th>
<th>Beta (α,β) prior distribution for sensitivity</th>
<th>Beta (α,β) prior distribution for specificity‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGID</td>
<td>0.2</td>
<td>0.4†</td>
<td>(4.46, 14.84)</td>
<td>451, 1</td>
</tr>
<tr>
<td>ELISA</td>
<td>0.25</td>
<td>0.5*</td>
<td>(3.88, 9.63)</td>
<td>449, 3</td>
</tr>
<tr>
<td>Skin test</td>
<td>0.7</td>
<td>0.2‡</td>
<td>(2.25, 1.52)</td>
<td>446, 6</td>
</tr>
</tbody>
</table>

* Expert was 95% sure that the sensitivity was less than this value
† Expert was 95% sure that the sensitivity was greater than this value
‡ Beta distributions for specificity for each test were constructed using the results from non-infected flocks: α = number of test-negative results + 1; β = number of test-positive results + 1

Table III. Skin test, AGID, and ELISA test positive results from 9 flocks (2015 sheep) that had sheep infected with MAP stratified against sheep age.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Skin test No. positive (%)</th>
<th>AGID No. positive (%)</th>
<th>ELISA No. positive (%)</th>
<th>No. sheep tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>99 (20.7)</td>
<td>7 (1.5)</td>
<td>8 (1.8)</td>
<td>478</td>
</tr>
<tr>
<td>2</td>
<td>116 (30.9)</td>
<td>8 (2.1)</td>
<td>16 (4.3)</td>
<td>375</td>
</tr>
<tr>
<td>3</td>
<td>83 (30.4)</td>
<td>15 (5.5)</td>
<td>12 (4.4)</td>
<td>273</td>
</tr>
<tr>
<td>4+</td>
<td>282 (31.7)</td>
<td>38 (4.3)</td>
<td>34 (3.8)</td>
<td>889</td>
</tr>
<tr>
<td>Total</td>
<td>572 (28.4)</td>
<td>68 (3.8)</td>
<td>70 (3.5)</td>
<td>2015</td>
</tr>
</tbody>
</table>
Table IV. Prior and posterior median and 95% probability intervals (PI) for estimates of sensitivity and specificity of AGID, ELISA for 3 cutoffs (0.20, 0.15, 0.10) and skin tests for ovine Johne’s disease

<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prior median</td>
<td>Prior 95% PI</td>
</tr>
<tr>
<td>AGID</td>
<td>0.222</td>
<td>0.077, 0.437</td>
</tr>
<tr>
<td>ELISA-0.20</td>
<td>0.276</td>
<td>0.089, 0.544</td>
</tr>
<tr>
<td>ELISA-0.15</td>
<td>0.276</td>
<td>0.089, 0.544</td>
</tr>
<tr>
<td>ELISA-0.10</td>
<td>0.276</td>
<td>0.089, 0.544</td>
</tr>
<tr>
<td>Skin test</td>
<td>0.613</td>
<td>0.146, 0.955</td>
</tr>
</tbody>
</table>
Table V. Paired comparisons of AGID and ELISA results at three different ELISA cutoffs, 0.20 (approved sheep cutoff), 0.15, 0.10 (approved cattle cutoff) in 2015 sheep from 9 U.S. Midwest sheep flocks. These were MAP infected flocks but the infection status of tested sheep was not known.

<table>
<thead>
<tr>
<th></th>
<th>Cutoff 0.20 + Neg control</th>
<th>Cutoff 0.15 + Neg control</th>
<th>Cutoff 0.10 + Neg control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ELISA +</td>
<td>ELISA-</td>
<td>ELISA +</td>
<td>ELISA-</td>
</tr>
<tr>
<td>AGID+</td>
<td>21</td>
<td>47</td>
<td>22</td>
<td>46</td>
</tr>
<tr>
<td>AGID-</td>
<td>49</td>
<td>1898</td>
<td>75</td>
<td>1872</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>1945</td>
<td>97</td>
<td>1918</td>
</tr>
</tbody>
</table>

Figure 1. Induration response (mm) of the 572 of 2015 sheep testing positive on the skin test from 9 flocks infected with MAP and the 5 sheep testing positive of 450 sheep from 5 non-infected flocks.
CHAPTER 4. EVALUATION OF THE GAMMA INTERFERON ELISA IN SHEEP SUBCLINICALLY INFECTED WITH MYCOBACTERIUM AVIUM SUBSPECIES PARATUBERCULOSIS USING A WHOLE CELL SONICATE OR A JOHNIN PURIFIED PROTEIN DERIVATIVE

A paper published in

Journal of Veterinary Diagnostic Investigation*

Suelee Robbe-Austermana, Judith R. Stabela, Mitchell V. Palmera

4.1 Abstract

The aim of this study was to estimate the sensitivity and specificity of the gamma interferon (IFN-γ) ELISA for paratuberculosis in sheep using receiver-operating characteristic analysis. Bacteriologic culture of tissues was used to define the reference positive population (n = 33). Two different reference negative populations were used, culture negative sheep from infected flocks (n = 77) and sheep from non-infected flocks (n = 358). This study also evaluated the accuracy of 2 different Mycobacterium avium subspecies paratuberculosis (MAP) antigen preparations, a whole cell sonicate (MpS) and a johnin purified protein derivative (PPD). The source of the reference negative

* Reprinted with permission of Journal of Veterinary Diagnostic Investigation, 2006, 18:189-194

aNational Animal Disease Center, Agricultural Research Service, USDA, Ames IA
sheep used in the analysis affected overall accuracy of the IFN-γ ELISA. The area under the curve was 0.683 (95% confidence interval 0.574-0.787) using culture negative sheep from infected flocks and 0.831 (0.764-0.889) using sheep from non-infected flocks for the MpS and was 0.809 (0.726 - 0.881) and 0.897 (0.862-0.925) for the johnin PPD. Using the MpS, the cut point that classified the most sheep correctly was an optical density reading of 0.20 for a sensitivity of 40.7% (19.4 - 57.6) and a specificity of 88.7% (77.0-95.7) or 97.6% (93.04-99.5) depending on the reference negative population used. Using the johnin PPD, the cut point that classified the most sheep correctly was 0.25 for a sensitivity of 66.7% (47.2-82.7) and a specificity of 93.5% (85.5-97.9) or 98.3% (96.4-99.4). The johnin PPD was more accurate at identifying MAP infected sheep than the MpS (p value = 0.034).

4.2 Introduction

Paratuberculosis or Johne’s disease, caused by *Mycobacterium avium* subspecies *paratuberculosis* (MAP), is a significant disease in ruminants worldwide. Current diagnostic tests (absorbed ELISA for antibody detection and fecal culture) typically detect animals only in the later stages of the disease after they start shedding the organism. For eradication of paratuberculosis from a flock or herd, however, it is critical for diagnostic tests to identify animals before they begin shedding the microorganism. Therefore, there is a need for alternative diagnostic tests that detect animals in the early stages of the disease.
It is thought that in the early stages of paratuberculosis the cell mediated immune (CMI) response predominates, and then wanes with advancing disease. Two CMI diagnostic tests for paratuberculosis have been proposed for field use, the skin test which measures delayed type hypersensitivity reaction in vivo, and the gamma interferon (IFN-γ) ELISA which measures IFN-γ production in response to antigen exposure in vitro.

The skin test has a long history in the literature. The avian tuberculin test was first proposed as a potential diagnostic test for paratuberculosis by Male in 1911, and again by M’Fadyean in 1916, using johnin as the antigen. Since then, over 100 papers have been published on the use of skin testing as a diagnostics tool, largely between 1911 though the 1960’s. The skin test gradually fell out of favor after repeated reports of failure to correlate test results with presence of the disease.

A diagnostic test using IFN-γ production to detect paratuberculosis was first proposed in 1992 and again in 1996. Since then, a commercial IFN-γ ELISA has been approved in the USA as an adjunct test for tuberculosis in cattle. IFN-γ ELISA’s have also been approved or are under consideration for approval for primates, humans and cervids. Although IFN-γ ELISA’s have been evaluated for paratuberculosis in cattle, sheep, and goats, no formal analysis of both sensitivity and specificity in naturally infected sheep has been done.

One advantage of the IFN-γ ELISA over the skin test is the ability to objectively measure the magnitude of the IFN-γ response by comparison with standardized positive and negative controls. The ELISA optical density (OD) readings can be used in a receiver-operating characteristic (ROC) analysis. ROC analysis allows for an overall
estimate of test accuracy (area under the curve) and provides the ability to optimize cut points depending on sensitivity and specificity needs.⁸

A particular problem with the evaluation of CMI diagnostic tests for paratuberculosis is the lack of a reference test for animals in the early stages of infection. Researchers are faced with attempting to validate diagnostic tests without an accurate method to confirm or deny positive test results. For paratuberculosis, bacterial culture from tissues is often used as a reference test. However, it has been suggested by other researchers that up to 100 tissues may need to be cultured to confirm infection in cattle.²⁸ The lack of a sensitive reference test confounds estimates of a potential diagnostic test’s accuracy. This is particularly true of tests capable of detecting early stages of MAP infection since tissue colonization would be minimal at that stage. With this concern in mind, the aim of this study was to evaluate IFN-γ ELISA using ROC analysis with bacterial culture positive sheep as the reference positive population, but with two different reference negative populations: culture negative sheep from infected flocks and sheep from non-infected flocks. This study also evaluated the accuracy of 2 different antigen preparations, a MAP whole cell sonicate (MpS) and a johnin purified protein derivative (PPDj).

4.3 Materials and Methods

4.31 Sheep

One hundred and ten sheep aged between 9 months and 5 years with no visible signs of paratuberculosis were purchased from four sheep flocks infected with paratuberculosis. These flocks had a history of at least one confirmed case of
paratuberculosis diagnosed by histology and confirmed with culture and/or PCR. Three of the infected flocks were Suffolk, located in Iowa (2) and South Dakota (1), and the other was a Polypay flock located in Minnesota. Sheep were purchased over a 2 year period beginning in the summer of 2003.

Three hundred and fifty-eight sheep from two MAP non-infected flocks were also used for an additional non-infected reference population in the ROC analysis. They were commercial white face crossbred flocks in Iowa located 200 km apart. No outside ewes have been introduced into these flocks in over 15 years. All sheep 6 months old or older in the flocks were included in the study during the spring/summer of 2004. Tissues were not taken for culture from these sheep. Both of these flocks were screened with the skin test, antibody ELISA, and the agar gel immuno-diffusion test for paratuberculosis in 1999 and all test positive sheep (2 skin test positive and 2 ELISA positive sheep from a total of 276) were examined upon necropsy and no evidence of paratuberculosis was found either grossly or microscopically. Procedures for histological examination are described in detail below. Bacteriologic culture was not done. Since 1999, these flocks have continued to maintain high biosecurity practices and routine necropsies on debilitated sheep have not revealed paratuberculosis. Field testing and necropsy protocols were approved by the National Animal Disease Center Animal Care and Use committee.

4.32 Bacterial culture and histopathology

Sheep were euthanized with 40 mg/kg of pentobarbital and the following 8 tissues were taken for bacteriologic culture and microscopic examination: ileo-cecal valve and associated mesenteric lymph node, ileum, distal jejunum and associated mesenteric
lymph node, mid jejunum and associated mesenteric lymph node, and proximal jejunum. Approximately 2 g of each tissue was homogenized in 25 ml of 0.75% hexadecylpyridinium chloride solution using a stomacher for 1 minute and then and allowed to settle for 30 minutes. Ten milliliters of supernatant was removed from just above the sediment line, transferred to a sterile tube and incubated at 37°C for 3 hours. Samples were then centrifuged at 1296 x g for 20 min, the supernatant was discarded and the pellet was resuspended in 1 ml of antibiotic mixture (100 µg/ml vancomycin, 50 µg/ml amphotericin B and 100 µg/ml of nalidixic acid in 0.9% brain heart infusion broth) and allowed to incubate overnight at 37°C. Then 200 µl was inoculated into radiometric liquid media vials supplemented with 1 ml of egg yolk, 600 µl of H2O, 5 µg of mycobactin J, and 100 µl PANTA PLUS and incubated at 37°C. Bottles were examined once a week for 12 weeks and bottles with a growth index ≥14 were confirmed with acid fast staining and by IS900 PCR using a technique described elsewhere. Samples were considered positive only if they were IS900 positive.

For histopathologic analysis, the tissues were routinely processed and paraffin embedded. Sections 5µm thick were stained with hematoxylin and eosin and examined by light microscopy. Adjacent sections were stained by the Ziehl-Neelsen technique to visualize acid-fast bacteria. Lesions were graded on a scale 1-3 defined as follows: 1, granulomatous lesions confined to lymphoid tissue and no acid fast bacteria (AFB); 2, granulomatous lesions in lymphoid tissue and in mucosa associated with Peyer’s patches with occasional AFB; 3, disseminated granulomas. Type 3 was further divided into subtypes a, b and c: type 3a, diffuse granulomas without modification of the cellular morphology and usually visible mycobacteria; type 3b, diffuse lepromatous granulomas
with many visible mycobacteria; type 3c, tuberculoid granulomas with few to no mycobacteria. The pathologist examining the sections was blinded to the culture and IFN-γ ELISA results of the sheep.

4.33 Sample collection and IFN-γ ELISA

For sheep in MAP infected flocks, samples for the IFN-γ ELISA were taken within 6 months of necropsy. Whole blood was collected into 10 ml heparinized blood collection tubes, kept at ambient room temperature and processed within 8 hours. One milliliter of blood was pipetted into separate wells on a sterile, 24-well tissue culture plate for each of the following treatments: 1) non stimulated negative control, 2) staphylococcal enterotoxin A (SEA), (positive control, 1 µg), 3) 10 µg johnin PPD (PPDj) made from MAP strain 19698 and, 4) 10 µg whole cell sonicate (MpS) also made from MAP strain 19698. Procedures for making PPDj and MpS are described in detail elsewhere. Due to a lack of MpS antigen, only 80 of 110 sheep from the MAP infected flocks and 123 sheep of the 358 sheep from the MAP non-infected flocks were tested with (MpS). To reduce bias, systematic randomization was used to select the sheep tested with MpS. The tissue culture plates were incubated for 18-24 hours at 37°C in 5% CO₂ humidified atmosphere. Plates were then centrifuged at 1296 x g for 5 minutes, and 500 µl of plasma was removed and stored at -80°C until assayed. A commercial IFN-γ ELISA kit was used according to the manufacturer’s instructions, with each sample run in duplicate. Results were read at 450 nm and the OD readings from the duplicate wells were averaged.
4.34 Statistical Analysis

Background levels of IFN-γ as measured by the non-stimulated wells were subtracted from the simulated wells within each animal. For results to be included in the analysis, the OD value of the positive control SEA minus the OD of the non-stimulated negative control must be $\geq 0.10$. ROC analysis and ROC comparison was calculated with Stata version 8\(^6\) using a nonparametric method.\(^6\) The reference positive population (bacterial culture positive sheep) and the two different reference negative populations (culture negative from infected flocks, sheep from non-infected flocks) were analyzed separately using the MpS antigen and then again using PPDj antigen. Several cut points were derived from each curve and exact binomial confidence intervals for sensitivity and specificity were calculated separately.\(^3\)

In order to compare the accuracy of the antigens, only the sheep that had results from both antigens were considered. ROC curves were recalculated and an algorithm was used to test the equality of the area under the curves.\(^6\)

4.4 Results

4.41 Bacterial culture and histopathology

Thirty-three (30\%) of the 110 sheep from the MAP infected populations were culture positive. The majority of those sheep (58\%) were either histologically negative or classified as type 1. Disseminated disease, with large numbers of AFB (type 3a) was only seen in 5 of the 33 sheep (Table 1).
4.42 ROC analysis

Changing the reference test negative sheep used in the ROC analysis affected the overall accuracy of the IFN-γ ELISA. The MpS antigen had an area under the curve (AUC) of 0.683 if bacterial culture alone was used as the reference test (Fig. 1A). However, using the positive bacterial culture sheep as the reference positive and sheep from non-infected flocks as the reference negative, improved the AUC to 0.830 (Fig. 1B). Similar improvements were seen by using the PPDj. The AUC was 0.809 (Fig. 1C) using only bacterial culture and 0.897 using the non-infected flocks (Fig. 1D). This improvement was seen because there were significantly more culture negative sheep in infected flocks with high IFN-γ ELISA readings than there were in the non-infected flocks. This consequently affected the specificity (Sp) estimates when a cut point was applied but not the sensitivity (Se) estimates (Table 2). The cut point that offered the best overall accuracy or classified the highest number of sheep correctly was 0.20 for MpS (Se = 40.7% (95% CI 19.4 - 57.6%) and Sp = 88.7% (77.0-95.7) for culture negative or Sp = 97.6% (93.04-99.5) for non-infected flocks). For the PPDj, the optimal cut point was 0.25 (Se = 66.7% (47.2-82.7) and Sp = 93.5% (85.5-97.9) or Sp = 98.3% (96.4-99.4)).

In order to determine which antigen was more accurate, only the sheep tested with both antigens were analyzed. Sheep from non-infected flocks were used for the reference negative population. The PPDj was more accurate than the MpS antigen at identifying MAP infected sheep: AUC of 0.898 vs. 0.811, p value = 0.034 (Fig. 2).
4.5 Discussion

The goal when evaluating diagnostic tests for diagnostic sensitivity and specificity should be to follow the World Organization for Animal Health (OIE) guidelines.\textsuperscript{18} While the number of animals used in this study falls short of these recommended guidelines, the data was sufficient to evaluate differences in the antigens and reference test populations used in this study. However, low animal numbers can mask important differences in diagnostic tests, antigens or intermittent cross reactivity problems and more work would need to be done address these important issues.

Others have investigated the IFN-\(\gamma\) ELISA specificity in cattle more thoroughly than this study did in sheep and found that specificity could be quite variable (71-99\%) between herds, batches of antigen, and the methodology used to establish a cut point.\textsuperscript{12} No widespread evaluation of specificity for the IFN-\(\gamma\) ELISA in sheep has been done. This study was only able to include 2 non-infected flocks, as it is particularly difficult to know the true flock status of paratuberculosis with the management practices in the USA. A more thorough evaluation of non-infected flocks is needed to confirm the relatively high level of specificity found in this study.

None of the 110 sheep purchased for this study from the MAP infected flocks displayed any signs of paratuberculosis. The majority of the culture positive sheep were also histologically negative or classified as type 1, clearly in the early stages of infection. Other studies have shown that fecal culture is 17\%-48\% sensitive in sheep with paucibacillary lesions and would likely be lower in sheep without histological lesions.\textsuperscript{29,30} If the IFN-\(\gamma\) ELISA can accurately classify these sheep in the early stages of infection, it would fulfill a need that is not currently being met by fecal culture and serology.
The overall accuracy of the IFN-γ ELISA in this ROC analysis was affected by the definition of the reference negative population. The specificity of the IFN-γ ELISA was lower (sometimes much lower depending on the cut point chosen) if culture negative sheep from infected flocks were used in the analysis. This brings up an interesting and important clinical question. Are those culture negative sheep in infected flocks that have a high IFN-γ response a potential threat or of any biological relevance? If so, bacteriologic culture of tissues may not be an adequate reference test for evaluating CMI based diagnostic tests.

There are three possible explanations for the difference in specificity depending on the reference negative population used. First, the sheep could have had an active infection, but bacterial culture was not sensitive enough to detect them. Second, the sheep could have been infected, but cleared the infection. Third, the sheep could have been exposed to MAP, but were not infected. It is unknown whether natural exposure without infection can cause a measurable CMI immune response. It is also not known how long an animal will have a measurable CMI response after it clears a MAP infection or more importantly, how often animals actually clear MAP infections. Understanding these events will be important for proper interpretation of positive results from the INF-γ ELISA. It may be possible that tests capable of detecting MAP in such early stages of infection will not very good at predicting clinical disease or even predicting shedding.

Current diagnostic tests (antibody ELISA and fecal culture) lack sensitivity in both sheep and cattle when they are in the early stages of MAP infection. Because of this, these tests are useful at detecting clinical animals, but they have little impact on negative predictive value. In other words, a negative test result does not greatly increase our
confidence the animal is not infected. This also hold true at the flock/herd level. Using pooled fecal culture one study demonstrated that 7 pools of 50 animals each would be required to achieve a 95% confidence level the flocks had <2% fecal shedders.\textsuperscript{22} Serology was unable to obtain that level of confidence even with whole flock testing. Although the studies would need to be done, the data in this paper suggests that it might be possible to increase the INF-\(\gamma\) ELISA cut point high enough to have the required specificity for a herd test and yet still maintain enough sensitivity to detect most \textit{MAP} infected flocks, regardless of the number animals shedding \textit{MAP} in their feces.

In this study, it appears that the johnin PPD is a better antigen than MpS. This is consistent with the findings from earlier literature when researchers were working on improving the sensitivity and specificity of the heat concentrated culture filtrates.\textsuperscript{7,17} Indeed, previous studies have demonstrated that the process of purifying johnins by precipitation with trichloroacetic acid improves the potency as well as specificity. This is likely due to the elimination of carbohydrates and glycolipids. It may well be the MpS could be improved by trichloroacetic acid precipitation. Whatever the case, this study suggests that the choice of antigen for stimulation of a CMI response is important. While PPD\textsubscript{j} yielded higher specificity and sensitivity than MpS in this study, there may be antigens superior to PPD\textsubscript{j} that have yet to be evaluated. Furthermore, antigen choice may need to be optimized according to species and stage of infection.

\subsection*{4.6 Acknowledgements.}

The authors would like to thank Marcia Hart and Megan Parlett for their technical assistance.
4.7 Sources and manufacturers

a Bovigam, Pfizer Animal Health, New York, NY

b Sigma-Aldrich; St. Louis, MO

c BACTEC 12B; Becton Dickinson, Sparks, MD

d Allied Monitor Inc., Fayette, MO

e PANTA PLUS; Becton Dickinson, Sparks, MD

f Falcon Multiwell; Becton Dickinson Labware, Franklin Lakes, NJ

g SEA, Toxin Technology, Sarasota, FL

h Johnin PPD 0201 lot 3, NVSL, Ames, IA

i National Animal Disease Center, Ames, IA

j Stata Corp, College Station, TX

4.8 References


Table 1. MAP culture and histology results from 110 sheep purchased for the study.

<table>
<thead>
<tr>
<th>Histology score</th>
<th>Culture Positive (%)</th>
<th>Culture Negative (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>12 (36)</td>
<td>77 (100)</td>
</tr>
<tr>
<td>Type 1</td>
<td>7 (21)</td>
<td>0</td>
</tr>
<tr>
<td>Type 2</td>
<td>5 (15)</td>
<td>0</td>
</tr>
<tr>
<td>Type 3a</td>
<td>5 (15)</td>
<td>0</td>
</tr>
<tr>
<td>Type 3b</td>
<td>2 (6)</td>
<td>0</td>
</tr>
<tr>
<td>Type 3c</td>
<td>2 (6)</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>33</strong></td>
<td>77</td>
</tr>
</tbody>
</table>

*a* Sheep were from 4 different flocks and did not show any clinical signs of MAP infection.
Table 2. Cut points for the IFN-γ ELISA using 4 different ROC analyses with 2 different antigens (MpS and PPDj) and two different reference tests (Culture positive and culture negative; and then culture positive and non-infected flocks).

<table>
<thead>
<tr>
<th>Cut point</th>
<th>Sensitivity&lt;sup&gt;a&lt;/sup&gt; (%)</th>
<th>95% CI (%)</th>
<th>Reference test: Culture + and Culture -</th>
<th>Reference test: Culture + and non-infected flocks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Specificity (%)</td>
<td>95% CI (%)</td>
<td>Specificity (%)</td>
</tr>
<tr>
<td>MpS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;= 0.05</td>
<td>70.37</td>
<td>49.82 - 86.25</td>
<td>49.06</td>
<td>35.06 - 63.16</td>
</tr>
<tr>
<td>&gt;= 0.10</td>
<td>59.26</td>
<td>38.80 - 77.61</td>
<td>69.81</td>
<td>55.66 - 81.66</td>
</tr>
<tr>
<td>&gt;= 0.15</td>
<td>51.85</td>
<td>31.95 - 71.33</td>
<td>81.13</td>
<td>68.03 - 90.56</td>
</tr>
<tr>
<td>&gt;= 0.20*</td>
<td>40.74</td>
<td>19.40 - 57.63</td>
<td>88.68</td>
<td>76.97 - 95.73</td>
</tr>
<tr>
<td>&gt;= 0.25</td>
<td>33.33</td>
<td>16.52 - 53.96</td>
<td>90.57</td>
<td>79.34 - 96.87</td>
</tr>
<tr>
<td>&gt;= 0.40</td>
<td>25.93</td>
<td>11.11 - 46.28</td>
<td>96.23</td>
<td>87.02 - 99.54</td>
</tr>
<tr>
<td>PPDj</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;= 0.05</td>
<td>80.00</td>
<td>61.43 - 92.29</td>
<td>55.84</td>
<td>44.07 - 67.16</td>
</tr>
<tr>
<td>&gt;= 0.10</td>
<td>80.00</td>
<td>61.43 - 92.29</td>
<td>75.32</td>
<td>64.18 - 84.44</td>
</tr>
<tr>
<td>&gt;= 0.15</td>
<td>76.67</td>
<td>54.11 - 87.72</td>
<td>81.82</td>
<td>71.38 - 89.69</td>
</tr>
<tr>
<td>&gt;= 0.20</td>
<td>70.00</td>
<td>50.60 - 85.27</td>
<td>84.42</td>
<td>74.36 - 91.68</td>
</tr>
<tr>
<td>&gt;= 0.25*</td>
<td>66.67</td>
<td>47.19 - 82.71</td>
<td>93.51</td>
<td>85.49 - 97.86</td>
</tr>
<tr>
<td>&gt;= 0.40</td>
<td>50.00</td>
<td>31.30 - 68.70</td>
<td>96.10</td>
<td>89.03 - 99.19</td>
</tr>
</tbody>
</table>

* Cut point with maximum accuracy (classified the highest number of sheep correctly)

<sup>a</sup> Sensitivity results were not affect by the reference negative population used in the ROC analysis.
Figure 1. ROC curves for IFN-γ ELISA in sheep. A) MpS antigen with the reference test being culture positive sheep and culture negative sheep from the same infected populations (n = 80); AUC = 0.683; 95% CI = 0.574 - 0.787. B) MpS antigen with the reference test being culture positive sheep and sheep from flocks without paratuberculosis (n = 153); AUC = 0.830; 95% CI = 0.764 - 0.889. C) PPDj antigen with the reference test being culture positive sheep and culture negative sheep from the same infected populations (n = 110); AUC = 0.809; 95% CI = 0.726 - 0.881. D) PPDj antigen with the reference test being culture positive sheep and sheep from flocks without paratuberculosis (n = 388); AUC = 0.897; 95% CI = 0.862-0.925.
Figure 2. ROC comparison of two antigens, MpS and PPDj in the IFN-γ ELISA with the reference test being culture positive sheep and sheep from flocks without paratuberculosis (n = 150); MpS AUC = 0.811; PPDj AUC = 0.898; p value = 0.034.
CHAPTER 5. TIME DELAY, TEMPERATURE EFFECTS AND ASSESSMENT OF POSITIVE CONTROLS ON WHOLE BLOOD FOR THE GAMMA INTERFERON ELISA TO DETECT PARATUBERCULOSIS

Reprinted with permission of

Journal of Veterinary Medicine B, Infectious Disease and Public Health*

Suelee Robbe-Austerman, DVM, MS, Adam C. Krull BS, Judith R. Stabel, PhD

5.1 Abstract

Our objective was to evaluate the effects of time and temperature on whole blood used in the gamma interferon (IFN-γ) ELISA for paratuberculosis along with evaluating 4 potential positive controls, and 4 different mycobacterial antigens for the ELISA. Nine adult Holstein cattle naturally infected with *Mycobacterium avium subspecies paratuberculosis* were used in a randomized complete block design. Forty-nine blood tubes were collected from each animal and held at 48.9, 37.8, 26.7, 21.1, 15.6, and 4.4°C for 0, 4, 8, 12, 18, 24, 32, 48, and 72 hours. Each blood tube was tested with 4 mycobacterial antigens (2 johnin PPD’s, an avian PPD and a whole cell sonicate) and 4 potential positive controls (concanavalin A (conA), phytohemagglutinin A (PHA),


*National Animal Disease Center, Agricultural research service, USDA, 2300 Dayton Ames, IA 50010

This work was supported by USDA ARS CRIS project # 3625-32000-062-00D
pokeweed mitogen (PWM) and Staphylococcus enterotoxin A (SEA)). After incubation for 24 hours, the plasma was assayed with a commercial IFN-γ ELISA. Blood stored at 21.1 and 15.6°C maintained the highest ELISA optical densities (OD) over time with severe reduction in OD values at or above 37.8°C. None of the potential positive controls exactly mimicked the antigen response. SEA and PWM were able to elicit a response after the whole blood quit responding to the antigen and conA underestimated the responsiveness. PHA was similar to the antigens on the average but there was significant disagreement among samples. The PPD’s were more potent at stimulating IFN-γ production than the whole cell sonicate.

Conclusion: Whole blood should be stored/transported at ambient room temperature and stimulated within 12 hours of collection.

5.2 Introduction

Paratuberculosis or Johne’s disease occurs throughout the world causing significant economic loss. Current serology and organism based diagnostic tests are insensitive and detect animals after they are infectious.(Stabel, 2000) Consequently, more sensitive tests are needed. Because a detectable cell mediated immune response (CMI) occurs early on during infection, diagnostic tests that use the CMI response to detect paratuberculosis have generated significant interest.

One such potential diagnostic test, the gamma interferon (IFN-γ) ELISA has already been approved in the United States as an adjunct test for detecting tuberculosis in cattle. The approval for tuberculosis has made this kit readily available and some diagnostic labs have begun to offer the test on a limited basis for paratuberculosis.
When an assay such as the IFN-γ ELISA requires the submission of whole blood with live lymphocytes that are healthy enough to respond to antigen stimulation, it becomes imperative to know how quickly the blood needs to arrive at the diagnostic lab and what effects hot or cold temperatures during shipping have on survival. Several studies have looked at time delay when using the IFN-γ ELISA in tuberculosis (Gormley, et al., 2004; Ryan, et al., 2000; Whelan, et al., 2004; Whipple, et al., 2001) and paratuberculosis (Jungersen, et al., 2002). All of these studies agree there was a reduction in IFN-γ production associated with time, but how much of a reduction was not always clear; some studies reported this reduction had little effect on the interpretation of the test, others disagreed. Most of these studies lacked sufficient sample numbers to discriminate a small to medium difference in sensitivity.

Less work has been done in regards to temperature effects. One study found cells shipped at 37°C had lower IFN-γ production than cells shipped at ambient room temperature. (Stabel and Whitlock, 2001)

The FDA has recently approved a IFN-γ ELISA for the diagnosis of latent tuberculosis in humans. (Mazurek and Villarino, 2003) For interpretation, this kit includes IFN-γ standards which allow the OD readings to be converted to international units (IU) of IFN-γ using a standard curve. A positive control, Phytohemagglutinin A (PHA) is also included. The results are reported as a ratio of test sample IU / PHA IU. PHA must also exceed 1.5 IU in order for the test results to be valid.

Although not promoted by the manufacturer of the bovine IFN-γ ELISA kit, many researchers have suggested using a mitogen or super antigen as a positive control. (Jungersen, et al., 2002; Manning, et al., 2003; Stabel, et al., 2003; Stabel and Whitlock, 2001) Two
primary reasons have been suggested. First, positive controls help identify anergic animals not capable of responding because of end stage paratuberculosis or possible tampering with immunosuppressive agents. The second is to ensure the lymphocytes in the whole blood have survived transport to the laboratory. A proper positive control would allow an animal that has a response to the mitogen, but not the antigen, be considered negative. An animal that failed to have a response to the mitogen should by definition also be negative to the antigen and an inconclusive result would be given.

Some researchers have documented that concanavalin A (conA) often fails to elicit a positive response when the lymphocytes are still able to generate a measurable response to an antigen. (Jungersen, et al., 2002) Other than this, there is little work published that actually validates any positive controls for the bovine IFN-\(\gamma\) ELISA for tuberculosis or paratuberculosis.

The aims of this study were to evaluate the effects of time delay and temperature on the IFN-\(\gamma\) response to paratuberculosis antigens, evaluate 3 mitogens and 1 super antigen as to their appropriateness as positive controls, and to assess the potency of 4 different mycobacterial antigens.

5.3 Materials and methods

5.3.1 Study design

A randomized complete block design was used. Nine naturally infected adult Holstein cattle (8 cows, 1 steer) that were part of the NADC Johne’s disease research herd and had a history of positive IFN-\(\gamma\) results were bled 2 times one week apart. For each bleeding, 30 10 ml heparinized blood collection tubes were filled. One tube from each animal was processed
immediately and the 9 other tubes were held at the prescribed temperature for the appropriate times. The tubes were held at temperatures 48.9, 37.8, 26.7 °C the first week and 21.1, 15.6, and 4.4 °C the next week. To evaluate time delay, tubes were held at their prescribed temperatures for 4, 8, 12, 18, 24, 36, 48, and 72 hours post bleeding.

5.32 IFN-γ ELISA

After the allotted time delay, the blood tube was gently mixed and 1 ml was added to 9 wells on a 24 well tissue culture plate. One well was left as a non-stimulated control with no additive. Other wells were treated with the following antigens: 10 µg Johnin PPD\(^a\) (PPDj3) made from MAP ATCC type strain 19698, 10 µg Johnin PPD\(^b\) (PPDj4) made from a recent cattle MAP isolate, 10 µg whole cell sonicate\(^c\) (MpS) made from strain 19698, and 10 µg aviun PPD\(^d\) (PPDav) (Pfizer, Omaha, NE). Four different positive controls were also used: 10 µg pokeweed mitogen\(^e\) (PWM), 1 µg staphylococcus enterotoxin A\(^e\) (SEA), 10 µg phytohemagglutinin A\(^d\) (PHA) and 10 µg concavalin A\(^d\) (conA).

All plates were incubated for 24 hours in 5% CO\(_2\) humidified atmosphere and then centrifuged at 1296 x g for 5 minutes. Approximately 500 µl of plasma were removed and stored at -80°C until assayed.

A commercial IFN-γ kit\(^f\) was used according to manufacturer’s instructions. Samples were run in duplicate and wells were read at 450 nm. As previously reported by others, there was a wide range of OD readings in the positive controls supplied by the kit for plates run at different times. This was most likely due to lab ambient temperature variations. The kit only required the positive control be > 0.7. In order to reduce plate to plate variation, the plates were standardized using the method suggested by others which was \(\text{OD}_C = [(\text{sample OD} - \text{OD}_{(c)})]/\text{OD}_{(c)}\).
NegC) \times (\text{mean Pos - mean Neg})/(\text{PosC - NegC})] + \text{mean Neg}, where OD_C is the calibrated OD, PosC and NegC are the positive and negative kit controls on the plate and mean Pos and mean Neg are the mean values of the positive and negative kit controls from all plates run. (Jungersen, et al., 2002) All results reported are calibrated OD of the test well – calibrated results of the non stimulated negative control.

Samples that had a non-stimulated well with a corrected OD reading of above 0.150 were eliminated from the analysis. A total of 5.2% of blood samples fit this criterion. These samples were seemingly randomly scattered throughout the times and temperatures, but one cow did contribute 60% of high non-stimulated samples.

5.33 Statistical analysis

Separate regression equations were calculated for each antigen and positive control treatment for OD response as a function of time and temperature. The mean OD responses of the 9 cows were used in the regression analyses. General linear model F-tests were used to determine if there were any overall equation differences between the 4 antigens, between the 4 positive controls, and between the overall antigens and positive controls. If the F-test showed that there were differences between the equations (treatments) tested, the 95% confidence intervals on the equation parameters for time and temperature were calculated and used for finding treatment comparison differences.

Repeated measures mixed model Analyses of Variance (ANOVA) were conducted to compare mean OD responses of the 4 antigens, the 4 positive controls, and between overall antigens vs. positive controls. Time and temperature variables were considered to be random effects in these models. If the ANOVA F-tests were significant at $p \leq 0.05$, then differences of
least squares means (LSMeans) was used as a multiple comparison test for determining mean differences among the 4 antigen means and among the 4 positive control means.

5.4 Results

All 9 cattle had IFN-γ ELISA responses to all mycobacterial antigens (mean OD = 1.51) and positive controls (mean OD = 2.01) at time 0. While there was some variation in OD values of the antigens initially (Range of 0.320-2.753), there were no significant differences between cattle at the rate (or slope) at which the cells lost the ability to produce IFN-γ. Consequently cattle which had lower IFN-γ ELISA responses initially became IFN-γ ELISA negative earlier (data not shown).

5.41 Antigen responses

In general, there was a loss of IFN-γ production that correlated with the time delay prior to stimulation. IFN-γ ELISA responses were maintained at levels comparable to time 0 for 4 hours except for the most extreme temperature (48.9°C) and then began to drop (Figure 1). Blood held at 48.9°C was unable to mount an IFN-γ ELISA response after 4 hours. Blood held at body temperature (37.8°C) for 8 hours lost 30% of the original OD value. The coldest temperature (4.4°C) evaluated was also detrimental, loosing 25% of the original OD value by 8 hours. Blood maintained around 15.6-21.1°C had the best overall ability to produce IFN-γ over time; loosing only 5% of the original OD value after 12 hours (Figure 2). Comparing the antigens, MpS was significantly less potent (p≤0.05) than all three PPD’s. PPDav tended to be lower than the johnin PPD’s although not significantly so (Figure 3). Only two different
times, 4 and 18 hours are shown in Figure 3, but the relationship holds for all the times except 72 hours where there is little response to any of the antigens.

5.42 Positive control responses

PHA was the positive control that mimicked the mycobacterial antigen responses most accurately across the range of times and temperatures tested. PWM and SEA were able to elicit strong IFN-γ responses after many antigen responses were not detectable. ConA lost its ability to elicit a response even when the antigens still consistently had detectable responses. Figure 4 demonstrates the typical relationship between an antigen (PPDj3) and the positive controls at 21.1°C over time. Similar patterns were found across all temperatures tested except for 48.9°C. Even though PHA was the most accurate positive control in terms of average response, the agreement at the individual cow level was variable; PHA would be able to stimulate a response but the antigen may not or vise versa.

5.43 Regression analysis

The regression equations used for predicting OD over time and temperature for each treatment were: \( Z = a + b [\ln(X)]^2 + c(Y)^3 \). Where \( Z = \) mean OD response, \( a = \) OD intercept, \( b = \) time parameter \( X = \) time, \( c = \) temperature parameter, and \( Y = \) temperature. All equations were statistically significant in explaining OD over time and temperature (\( p < .0001 \)). Parameter estimate slopes and intercept coefficients in the models were also significant (\( p \leq .0001 \)), indicating true, non-zero contributions to the prediction equation. The only exception was the temperature coefficient for ConA positive control, which was significant at \( p = .023 \). Table 1 contains the intercept, time and temperature parameters for the antigens
and table 2 for the positive controls. There were not enough data points at colder temperatures to accurately characterize the behavior of the blood stored below 15.6°C with the regression equations.

### 5.5 Discussion

The 9 cows chosen for this study were selected based on previous history of consistent positive IFN-γ ELISA responses. Consequently most of these cows had strong IFN-γ ELISA responses and this biased our results causing weak and moderately positive IFN-γ ELISA cows to be underrepresented.

Nonetheless, there was enough variation in initial OD readings to evaluate differences in rates of IFN-γ production loss over time and temperature between individual cows. The rates were remarkably similar regardless of initial OD results suggesting that rates of IFN-γ loss were independent of the strength of initial OD response. This important finding suggests that regression equations can be accurately used to predict the potential time delay and temperature effects on whole blood and how these time and temperature parameters may impact future test interpretations. It is important to note that even though the cow’s individual slopes did not vary, there were significant differences in slopes among the antigens. Care must be taken when extrapolating these equations and slopes to other antigens not used in this paper.

There was a significant difference in IFN-γ ELISA responses to antigens in this study. The PPD’s were more potent than the MpS antigen at stimulating IFN-γ. Similar results were found in sheep in a previous study.(Robbe-Austerman, et al., March, 2006) If more cattle had been tested, the johnin PPD’s would have likely also been statistically more potent than
PPDav. This is contrary to others who reported higher responses to PPDav than to Johnin PPD. (Jungersen, et al., 2002) There are several reports in the literature about the variability of potency and lack of lot to lot consistency in all the PPD’s. (Johnson, et al., 1949; Landi, et al., 1975) Because of this lack of consistency and cross reactivity with other mycobacteria, there is strong interest in developing specific antigens for use in the IFN-γ ELISA for paratuberculosis. In bovine tuberculosis IFN-γ assay the specific antigens, ESAT 6 and CFP 10, are not as potent as tuberculin PPD. (Waters, et al., 2004) This relative lack of potency of these specific proteins could increase the consequences of a delay in setting up blood samples for the IFN-γ ELISA. (Whelan, et al., 2004) It is likely similar consequences were noticed in the FDA approved human tuberculosis IFN-γ ELISA kit as samples are required to be set up with in 12 hours. (Mazurek and Villarino, 2003)

It is surprising that the majority of the positive controls used in the literature failed to mimic the IFN-γ production loss over time of the mycobacterial antigens. PWM and SEA are inadequate positive controls and give a false sense of security by maintaining a positive response long after the cells could no longer respond to the mycobacterial antigens. This study also confirms other work which reported that con A often failed to elicit a response when cells still had the ability to respond to the antigen. (Jungersen, et al., 2002). PHA on average mimicked the antigen curves most accurately. This is likely the reason why PHA is the positive control used in the human tuberculosis IFN-γ ELISA kit. However, PHA should be validated for each species tested as it fails to stimulate IFN-γ production in sheep whole blood.

The IFN-γ ELISA has potential as a diagnostic test for paratuberculosis as well as tuberculosis. However specimen handling is of utmost importance. To maximize the
accuracy of the test, whole blood should be stimulated within 12 hours of sampling, and the
storage/transport temperature prior to stimulation should be maintained between 15.6 and
21.1°C.

5.6 Footnotes

a National Veterinary Services Laboratory, USDA, Ames, IA

b National Animal Disease Center, ARS, USDA, Ames, IA

c Pfizer Animal Health, Omaha, NE

d Sigma Chemical, St. Louis, MO

e Toxin Technology, Sarasota, FL

f Bovigam, Pfizer Animal Health, Omaha, NE

g TableCurve 3D Version 3.01, SPSS Chicago, IL

h PROC MIXED SAS Version 9.1.3, SAS Institute, Cary, NC USA

5.7 References

effect of the tuberculin test and the consequences of a delay in blood culture on the
sensitivity of a gamma-interferon assay for the detection of Mycobacterium bovis

The relationship of the allergens of Mycobacterium paratuberculosis, Mycobacterium
tuberculosis var. avium, bovis, and hominus and Mycobacterium phlei. Am. J. Vet.
Res. 138-141.

Jungersen, G., A. Huda, J. J. Hansen, and P. Lind, 2002: Interpretation of the gamma
interferon test for diagnosis of subclinical paratuberculosis in cattle. Clin Diagn Lab
Immunol. 9, 453-460.

Dev Biol Stand. 29, 393-411.


Table 1. Regression equation parameters calculated for mycobacterial antigens used in the IFN-γ ELISA in 9 adult Holstein cattle with paratuberculosis.

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Intercept (OD at 0)</th>
<th>Rank</th>
<th>Time Parameter</th>
<th>Temp Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPDj4</td>
<td>2.533</td>
<td>1</td>
<td>-.12936 b</td>
<td>-1.813*10^-5 b</td>
</tr>
<tr>
<td>PPDj3</td>
<td>2.404</td>
<td>2</td>
<td>-.12743 b</td>
<td>-1.559*10^-5 b</td>
</tr>
<tr>
<td>PPDav</td>
<td>1.863</td>
<td>3</td>
<td>-.10169 b</td>
<td>-1.212*10^-5 ab</td>
</tr>
<tr>
<td>MpS</td>
<td>1.342</td>
<td>4</td>
<td>-.06558 a^1</td>
<td>-8.61*10^-6 a</td>
</tr>
</tbody>
</table>

^1 Slope parameter estimates followed by the same letter within a column are not significantly different based on overlap of the 95% confidence intervals.
Table 2. Regression equation parameters calculated for the positive controls used in the IFN-γ ELISA in 9 adult Holstein cattle with paratuberculosis.

<table>
<thead>
<tr>
<th>Positive Control</th>
<th>Intercept (OD at time 0)</th>
<th>Time Rank</th>
<th>Time Parameter</th>
<th>Temp Rank</th>
<th>Temp Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEA</td>
<td>3.367</td>
<td>1</td>
<td>-.15144 c</td>
<td>2</td>
<td>-2.027*10^{-5} bc</td>
</tr>
<tr>
<td>PWM</td>
<td>3.996</td>
<td>2</td>
<td>-.14978 bc</td>
<td>1</td>
<td>-2.515*10^{-5} c</td>
</tr>
<tr>
<td>PHA</td>
<td>2.079</td>
<td>3</td>
<td>-.10456 b</td>
<td>3</td>
<td>-1.307*10^{-5} b</td>
</tr>
<tr>
<td>ConA</td>
<td>0.721</td>
<td>4</td>
<td>-.03816 a</td>
<td>4</td>
<td>-3.36*10^{-6} a</td>
</tr>
</tbody>
</table>

Slope parameter estimates followed by the same letter within a column are not significantly different based on overlap of the 95% confidence intervals.
Figure 1. Effects of time and temperature on the OD readings of the IFN-γ ELISA using the mean response of 4 antigens, PPDj3, PPDj4, PPDa and MpS, in 9 adult Holstein dairy cattle naturally infected with MAP.
Figure 2. Impact of temperature on the OD readings at each time point for the IFN-γ ELISA using mean responses of 4 antigens, PPDj3, PPDj4, PPDa and MpS, in 9 adult Holstein dairy cattle naturally infected with MAP.
Figure 3. Average IFN-γ ELISA OD responses of 9 adult Holstein cattle naturally infected with MAP to 4 mycobacterial antigens at a time delay of 4 hours and 18 hours and across 5 different temperatures.
Figure 4. The relationship of the 4 positive control OD responses to the PPDj3 OD response over time at 21.1°C storage temperature using the mean of 9 adult Holstein dairy cattle.
CHAPTER 6. SKIN TEST AND GAMMA INTERFERON ENZYME-LINKED IMMUNOSORBENT ASSAY RESULTS IN SHEEP EXPOSED TO DEAD MYCOBACTERIUM AVIUM SUBSPECIES PARATUBERCULOSIS ORGANISMS

A paper published in Journal of Veterinary Diagnostic Investigation

Suelee Robbe-Austerman, a b Judith R. Stabel, a Daniel G. Morrical c

6.1 Abstract

Cell mediated immunity (CMI) diagnostic tests, such as the gamma interferon enzyme-linked immunosorbent assay (IFN-γ ELISA) and the johnin skin test, have the potential to detect animals infected with Mycobacterium avium subspecies paratuberculosis (MAP) early in the course of the disease. While these CMI tests tend to be relatively specific in non-infected flocks, in MAP infected flocks these tests often identify animals that can not be confirmed infected by any other reference test including necropsy and culture. The aim of this study was to determine if antigen exposure by inhalation or oral ingestion of killed MAP organisms would cause a detectable CMI


a Bacterial Diseases of Livestock Research Unit; National Animal Disease Center

ARS/USDA, Ames, IA 50010

b Veterinary Diagnostic and Production Animal Medicine, Iowa State University, Ames, IA 50011

c Deptment of Animal Sciences, Iowa State University, Ames, IA, 50011
response in sheep. Forty-eight lambs 4 months of age were randomly divided into a control group, an orally exposed group (dosed with \(1 \times 10^{10}\) autoclaved MAP organisms three times) and an inhalation exposed group (dosed once with \(1 \times 10^5\) dead organisms). Lambs were skin tested and/or bled pre-exposure, and 1, 2, 3, 4, 12 months post exposure. No significant difference was seen with either the oral or inhalation exposed groups of lambs vs. controls with either the IFN-\(\gamma\) ELISA or the skin test at any time pre or post exposure. These results suggest that infection/invasion of MAP organism must occur in order to have a positive skin test or IFN-\(\gamma\) ELISA beyond the false positive rate. Simple exposure is not enough to elicit a detectable CMI response.

6.2 Introduction

*Mycobacterium avium* subspecies *paratuberculosis* (MAP) causes Johne’s disease which is one of the major diseases that cause chronic, severe weight loss and eventual death in sheep and other ruminants. Strong interest has been expressed in developing diagnostic tests which detect animals in the early stages of disease, as current diagnostic tests (fecal culture and serology) usually detect animals only after they start to shed the organism. Because the cell mediated immune (CMI) response predominates in the early stages of MAP infection, diagnostic tests based on CMI responses hold considerable promise.\(^9\) Two such tests, the johnin skin test and the gamma interferon (IFN-\(\gamma\)) ELISA have proven to be relatively specific (>98%) in sheep from non-infected flocks, provided a specific and sensitive antigen is used.\(^7,8\)
However when these CMI based diagnostic tests are used in flocks or herds with infected animals and a reference test such as histology or tissue culture is used, specificity drops to below 90%. Animals in these infected populations that have positive skin test responses or have high optical density (OD) IFN-γ ELISA readings but do not appear to be infected using the reference test are of particular interest and must be more clearly defined if these tests are to have clinical value in infected populations.

Three potential explanations for these CMI test positive but reference test negative sheep have been proposed in a previous paper: a) the reference test is not sensitive enough to identify all infected animals; b) the animals were infected in the past but have recovered and harbor a residual immune response; and c) animals were exposed to MAP antigens but infection did not occur. If the third hypothesis occurs, these tests will likely have limited use in infected populations due to environmental contamination. While all of these hypotheses need to be tested, the objective in this study was to evaluate the exposure-without-infection hypothesis. Specifically, the aim was to determine if antigen exposure by inhalation or oral ingestion of killed MAP organisms would cause a detectable CMI immune response.

6.3 Materials and Methods

A total of 48 white faced cross bred ewe lambs 4 months of age from the Iowa State University McNay research flock were randomly assigned to one of 3 groups. This flock of 300 ewes was managed as a typical Midwest commercial winter lambing flock and had been used as a non-infected flock in previous studies. The flock was rigorously screened for Johne’s disease in 1999-2001 as previously described. Since then, cull and thin animals have been screened via culture or necropsy and strict biosecurity practices have been in place. The
flock has been closed to outside ewes for over 15 years, and only skin test negative rams have been introduced.

The MAP strain used in this study was MAP ATCC strain 19698 (bovine strain). This strain was chosen because of ease of differentiating it from a wild type sheep strain in the event the flock became infected with MAP after the study started. It is also the strain routinely used in the authors’ laboratory to experimentally infect calves.\textsuperscript{11,12} The bacteria were grown in Middlebrook 7H9 liquid media (pH 6.0) supplemented with 10% oleic acid albumin dextrose complex \textsuperscript{a}, 0.05% Tween 80\textsuperscript{b}, and 2 mg/ml mycobactin J \textsuperscript{b}. MAP cultures were grown to log phase at an optical density 450 nm (OD\textsubscript{450}) of 0.4, at 37\textdegree C without shaking. Bacteria grown to an optical density of 0.4 have been determined to be at a concentration between 1 \times 10\textsuperscript{7} and 1 \times 10\textsuperscript{8} by plating out serial 10-fold dilutions on Herrold’s egg yolk media\textsuperscript{a} (HEYM) and adjusted to obtain a final concentration of 1 \times 10\textsuperscript{10} for the oral drench or 1 \times 10\textsuperscript{5} for the aerosolization. The bacteria were then autoclaved at 120\textdegree C at 17 pounds per square inch for 30 min and then inoculated on to HEYM and monitored for growth for 8 weeks.

The first group of lambs was orally drenched 3 times one week apart with 1 \times 10\textsuperscript{10} of killed MAP organisms for a total of 3 \times 10\textsuperscript{10}. Group 2 lambs were forced to inhale 1 \times 10\textsuperscript{5} one time using a nebulizer and a mask as described elsewhere.\textsuperscript{5} The last group was left as controls. The ewe lamb groups were housed separately in 20 m\textsuperscript{2} pens in an open sided barn until one week after the final oral dose was given and then the animals were housed together in a small 3 acre pasture. Lambs were skin tested using 0.1 ml of a 1 mg/ml purified protein derivative (PPDj)\textsuperscript{c} in the axillary region and palpated for induration 72 hours later as described previously\textsuperscript{7} and blood was collected in 10 ml heparinized blood tubes immediately
before exposure, and then bled once every 30 days and skin tested every 60 days for 4 months and then at 12 months. Protocols were approved by Iowa State University’s Institutional Animal Care and Use Committee.

Whole blood for the IFN-γ ELISA was transported at ambient room temperature and stimulated within 8 hours as described previously. Briefly, the whole blood was allocated into 3, 1 ml aliquots. One well was left as a negative control, one well was stimulated with 1 µg of staphylococcus enterotoxin A (SEA) as a positive control and one well was stimulated with 10 µg of PPDj. The samples were incubated at 37°C for 24 hours and the plasma harvested. A commercial IFN-γ ELISA kit was used according to manufacturer’s instructions. Samples were considered valid if the optical density (OD) of the SEA well was > than 0.10. Results were reported subtracting the unstimulated well OD values from the wells stimulated with purified protein derivative (PPDj). The same PPDj that was used in the IFN-γ ELISA was used in the skin test.

For statistical analysis, IFN-γ ELISA results from the 3 groups were compared at each time point using ANOVA in Stata. The categorical skin test data were analyzed using the Fisher’s exact test.

### 6.4 Results

There were no differences in either the mean OD value of the IFN-γ ELISA or the skin test between the groups of lambs, either before or after exposure to MAP organisms. (Table 1). If a cut point at or above 0.2 was used in the IFN-γ, no lambs would have tested positive at any time throughout the experiment. A previous ROC analysis found a cut point
of .25 using this particular PPDj in the IFN-γ ELISA was determined to have the best overall accuracy.8

6.5 Discussion

Due to the expense and lack of required facilities a positive control group (sheep dosed with live organisms) was not used. While this was a significant study design flaw, other researchers have found that live oral inoculations in sheep of laboratory propagated MAP organisms reliably produced CMI responses at lower doses than used in this study before 4 months post infection.1,2,6 The authors were not aware of any MAP infection studies using a nebulizer, so the inhalation dose 1 x 10^5 was similar to the high dose regimen used in deer with Mycobacterium bovis using the same apparatus, and it is unlikely animals would naturally inhale a higher number of dead organisms.5 A bovine type strain was used in this study rather than a sheep strain. While there are some questions about the differences in strain pathogenicity, experimentally both strains cause similar early CMI responses when live organisms were used.10

The authors have shown in previous papers that the skin test and IFN-γ ELISA can be specific in MAP non-infected populations. This has not been true using post mortem testing of animals residing in infected populations as the reference test. While it is likely the reference tests lack sensitivity, animals in non-infected populations have no opportunity for environmental exposure to antigens of dead organisms. Because there is interest in using these tests in infected populations, the potential causes of positive CMI test results must be understood. The results of this current study suggest that in order for a detectable CMI response as measured by the skin test or the IFN-γ ELISA to occur against MAP, actual
infection/invasion with MAP organisms rather than simple exposure must occur. Therefore, the assumption can be made that positive test results (beyond the false positive rate) would be the result of infection, not just simple exposure to dead organisms and they are not identified as infected at necropsy because the reference tests are not sensitive enough or the sheep are no longer infected.

What is not known is if, and how often animals recover from an infection with MAP organisms? If most animals that get infected become latently infected and harbor the organism, those animals are likely to be potential threats to the rest of the flock or herd. A producer could then potentially use CMI test results as a selection tool for replacement animals. On the other hand, if most infected animals recover, a single positive CMI test result may have little meaning. Indeed, it seems reasonable that in heavily contaminated environments that there can be nearly 100% infection rates, some animals may eliminate the organism, some may remain latently infected, and a small number of those may come down with clinical Johne’s disease. Whatever the case, further understanding the meaning of a positive CMI result is certainly needed for proper positive test interpretation. This would likely require longitudinal studies.

6.6 Sources and manufactures

a Becton Dickinson Microbiology, Sparks, MD.

b Allied Monitor Inc., Fayette, MO.

c Johnin PPD 0201 lot 3, NVSL, Ames, IA.

d SEA, Toxin Technology, Sarasota, FL.

e Bovigam, Prionics, Omaha, NE, USA.
6.7 References

Table 1. Average gamma interferon ELISA OD values and number of animals with positive skin test results in lambs pre and post exposure to autoclaved *Mycobacterium avium* subspecies *paratuberculosis*.

<table>
<thead>
<tr>
<th></th>
<th>Pre exposure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IFN-γ (St dev)</td>
<td>Skin test</td>
<td>IFN-γ (St dev)</td>
<td>Skin test</td>
<td>IFN-γ (St dev)</td>
<td>Skin test</td>
</tr>
<tr>
<td>Control (n=16)</td>
<td>0.015 (0.035)</td>
<td>0</td>
<td>0.022 (0.051)</td>
<td>0.020 (0.035)</td>
<td>1</td>
<td>0.019 (0.036)</td>
</tr>
<tr>
<td>Inhalation (n=16)</td>
<td>0.008 (0.029)</td>
<td>1</td>
<td>0.023 (0.065)</td>
<td>0.019 (0.034)</td>
<td>0</td>
<td>0.014 (0.052)</td>
</tr>
<tr>
<td>Oral (n=15)</td>
<td>0.020 (0.027)</td>
<td>0</td>
<td>0.017 (0.041)</td>
<td>0.016 (0.024)</td>
<td>0</td>
<td>0.008 (0.038)</td>
</tr>
</tbody>
</table>
CHAPTER 7. GENERAL SUMMARY

7.1 General Discussion

CMI-based diagnostic tests (skin test and IV + temperature monitoring test) were the first diagnostic tests available for paratuberculosis. But after years of testing and attempts at improving the antigens, researchers eventually rejected CMI tests for antibody and organism detection tests which correlated better to clinical disease. Unfortunately, as study designs improved and more rigorous reference tests such as bacterial culture of tissues were used, antibody and organism detection tests have not been shown to be as sensitive as researchers had hoped.

Even so, fecal culture and the antibody ELISA have enjoyed wide use in the field over the last 15 years. However, producers have been frustrated with inability of these tests to detect subclinical animals. Producers want diagnostic tests that can be accurately used prior to purchase or prior to selecting replacement animals.

The need for diagnostic tests that detect MAP infections earlier in the course of the disease was the impetus for the research in this thesis. It was logical to consider that if CMI based diagnostic tests can be useful at diagnosing subclinical *Mycobacterium bovis* infections, they should also be useful at diagnosing subclinical *Mycobacterium avium* subspecies *paratuberculosis* infections, especially as research has shown the immune response against both organisms is similar. However, this was the line of reasoning that kept so many researchers pursuing the skin test for paratuberculosis in the 1930’s through the 1950’s; therefore, if researchers failed to develop useful CMI tests for
paratuberculosis in the past, it is critical to understand the problems they encountered and avoid or address those problems directly.

First, they were unable to produce a PPD that was specific in experimental infections to MAP and not \textit{M. avium}. This caused a great deal of concern, especially as \textit{M. avium} was often endemic in local farmers’ poultry flocks. Unlike with \textit{M. bovis} infections, there was not a differential response with johnin PPD and avian PPD when cattle were experimentally infected with \textit{M avium} or MAP. While this was still a concern, and the johnin PPD’s in the studies reported in this thesis are likely just as cross-reactive, most livestock operations in the US are not in close contact with poultry. Furthermore, flocks or cattle herds tested for these projects that were non-infected with MAP, had a fairly low false positive rate. See Appendix A for a detailed report of the specificities of different antigen products tested. A more thorough investigation of specificity throughout different regions of the country should be conducted to ensure false positive results remain low and are consistent in different environments. However, before this can be done a method for producing a consistent antigen needs to be in place.

The second problem researchers had previously encountered was the lack of a test animal to establish true infection status for paratuberculosis. For example, bovine tuberculosis had the guinea pig as the test animal. The guinea pig inoculation test was more sensitive than tissue culture or histology and provided researchers with a reasonable gold standard and confidence in a positive CMI result. Without a reliable gold standard for paratuberculosis, researchers were unable to determine accurate specificity estimates.

Because researchers in paratuberculosis did not have a reliable gold standard to evaluate infection status, researchers attempted to correlate CMI test results to clinical
Johne’s disease. Often the herds enrolled in trials evaluating CMI diagnostic tests were heavily infected herds with high levels of clinical disease. The skin test performed poorly in these situations, often failing to identify animals that were clinical or animals that would develop clinical disease in the near future. The skin test would also identify animals with no evidence of disease on necropsy.

Consequently, when it was decided to begin re-investigating the potential use of CMI diagnostic tests for this project, studies were designed to avoid the need for a perfect or near perfect reference test. In the first study a statistical approach (Bayesian 3 in 1 population model) was used that allowed for estimates of sensitivity and specificity without the use of a gold standard. This paper gave similar estimates on sensitivity and specificity that others have found with the single tuberculin skin test for *M. bovis*.

A follow-up to this study was a repeat visit to one of the farms 12 months later with another whole flock skin test. The results of this visit demonstrate the lack of repeatability of the skin test. There were 91 sheep that initially tested positive, 60 remained positive 12 months later. Thirty-four sheep that initially were negative on the first test became positive on the second test. A summary of these results can be found in Appendix B. This is consistent with the historical literature. Other researchers have called this a variation or natural flux in the allergic response.

The next study using the IFN-γ ELISA and ROC analysis showed similar sensitivity and specificity to the skin test. This study also showed that bacterial culture of tissues failed to identify many CMI positive animals in infected populations. This suggested that either, 1) bacterial culture of tissues was a poor gold standard, 2) previously infected animals had eliminated the infection and still had a residual CMI
response, and/or 3) animals were exposed to antigens in the environment that caused a
CMI response. The likely cause is probably a combination of the first 2 hypotheses.

The final study we conducted addressed the third hypothesis; environmental
exposure to large amounts of antigen would cause a CMI response. In this study we
forced animals to ingest or inhale large numbers of dead organisms. We did not detect
any response above the false positive rate in the control group. This suggests that
infection must occur in order to get a detectable CMI response.

7.2 Comparison of the IFN-γ ELISA and the skin test

Often the question is asked; which of the tests, the skin test or the IFN-γ ELISA,
is the better test? Both tests have their advantages and disadvantages. Both rely on the
CMI response, this adds an element of error that diagnostic tests relying on detecting the
organism directly do not have. Infected animals may not develop a detectable CMI
response, or animals may be sick or stressed, thus suppressing their immune system and
potentially giving a false negative result. There is also evidence that the CMI response is
variable throughout the incubation period of disease.

With regard to the specifics of the skin test and the gamma interferon, there are
advantages and disadvantages attributed directly to the tests themselves. The skin test can
be inexpensive, easy to perform in remote areas with very little equipment required. This
test relies on the delayed-type hypersensitivity reaction and requires multiple cytokines
and cells to be involved thus avoiding the potential hyper-reactivity of one cytokine. The
disadvantages of the skin test are; animals have to be handled twice; the skin
measurement is subjective and difficult to standardize; it can be difficult to identify the
exact area of injection; a bug bite or a localized infection due to a contaminated needle could accidentally be considered an induration. In Appendix A, indurations were occasionally attributed to animals receiving phosphate buffer solution or sham injection.

The advantages of the IFN-γ ELISA are: animals only have to be handled once for a blood draw; it is much easier to test multiple different antigens and less antigen is required; the IFN-γ response is measured objectively. The disadvantages of the IFN-γ ELISA are: it’s an expensive, time consuming and technically difficult test to run; samples have to be handled carefully and submitted to the lab within 12 hours; it relies on only one cytokine, IFN-γ to respond to the antigen. Appendix C details a flock with severe lesions in the mouth due to foxtail hay. The skin test only identified 3 sheep as positive, the IFN-γ ELISA was invalid in 9 of the 25 sheep tested due to high levels of background IFN-γ. Of the 16 sheep with normal background levels of IFN-γ, seven were positive with lot 3 johnin and five were positive with lot 4 johnin. Only 2 of these sheep were subsequently identified as infected with MAP, both of them were skin test positive.

7.3 Weakness of the studies in this thesis

The study described in detail in chapter 3 relied on a Bayesian 3 in 1 population model. A non-informative prior of (1,1) was used in a sensitivity analysis to look at the potential bias of the expert opinion offered by the author of this thesis. Even though the sensitivity analysis did not show that the priors affected the outcome, a more thorough sensitivity analysis should have been done using different informative priors. The author of the paper provided the expert opinion after she did the tests, and hence was likely biased when she provided the priors. The researchers were also not blinded as to the
potential infection status of the 14 flocks tested. Because reading of the skin test is a subjective measurement, it is possible prior knowledge may have biased the skin test readings. Furthermore, as large numbers of sheep are palpated for an induration with no positive response, it is possible for the skin test reader to expect negative results and not palpate subsequent animals as carefully. When conducting the analysis and writing the paper, the amount of dependence between the antibody ELISA and the AGID should have been included. Another program, TAGS (v. 2.0 at bayes.math.montana.edu) was initially considered (See Appendix D). Even though the results are similar between the models: point estimates for sensitivity for the ELISA, AGID and skin test were .083, 080, and .733 for the Bayesian model and .084, .083 and .700 for the TAGS program, there was enough correlation between the ELISA and AGID for TAGS to warn the model does not fit, and a model that allowed for dependence between 2 tests was needed.

The study described in chapter 4 used ROC analysis to compare 2 different antigens in the IFN-γ ELISA. The method used a non-parametric approach to calculate the curve and an algorithm based on DeLong’s paper in 1988 (see reference #6 on page 73) to compare the area under the curve (AUC) of the 2 tests. This method compares the entire AUC; however, if the shapes of the curves are visualized (see figure 2 on page 79) there is a point where the curves meet that may be in the region where sensitivity and specificity are maximized. If that is the case, then there would not be differences in the antigens when the tests were applied to clinical situations. A total of 150 sheep were used in the comparison with only 31 of those testing positive to the reference test. This is not a large enough sample size to definitively say that the difference in the curves at the
area of interest was or was not clinically significant. The study should be repeated with more sheep.

The study in Chapter 5 looks at the effect of time and temperature delays on whole blood used for the IFN-γ ELISA. Even though the paper states a randomized complete block design was used, this was not the case. There was no randomization involved. Every cow had a blood sample in every treatment.

In the statistical methods section it states that “A repeated measures mixed model Analyses of Variance (ANOVA) were conducted” and that “Time and temperature variables were considered to be random effects”. In fact, time and temperature were considered fixed effects, not random. Furthermore, confidence intervals were used to determine statistical significance rather than using a method that tests for significance directly. By using confidence intervals, some of the antigens and positive controls that were considered not significant may have been significant and were missed.

The regression equations that were used in the paper do not accurately reflect the data at lower temperatures. All 8 regression equations show that lower temperatures have a protective effect on whole blood. This was not the case. The likely problem is there were only 2 temperature points that were lower than the “ideal” temperature. Other higher order variables should have been considered when building this model. This may have improved the fit and better characterized the data at lower temperatures. Because of the unusual curvature of the data, (see Figure 2 on page 95) it is possible that one equation is not sufficient to describe the survivability of the whole blood at both low and high temperatures. An equation modeling the decline in survivability should have been
calculated to look at the temperatures below the optimum range, and then the equation in the paper can be used to model the effects that higher temperatures have on survivability.

The study detailed in Chapter 6 looks at CMI responses to dead organisms. In this study, 4 months post exposure there were 2 animals out of 15 that had positive skin test responses to the PPD. None of the 16 control sheep tested positive. A Fisher’s exact test was used to compare the 2 groups and while the results were not significant, the lack of significance may be due to a lack of power rather than artifact due to chance. Appendix A details the specificity of different PPD’s used in the papers included in this thesis. This particular paper used Johnin 0202 which has a point estimate specificity of 84%. Having 2 false positives out of 15 would not be unexpected. However, before a conclusion that the skin test is not affected by dead organisms, the trial should be repeated preferably with more animals. The raw data suggests a 13% higher response rate in the orally exposed sheep at 4 months post exposure. In order to determine if a 13% difference in response rate was significant using the Fisher’s exact test 30 animals per group should be used.

7.4 Preliminary Results of a longitudinal study evaluating the skin test and IFN-γ ELISA in calves

Interpreting a positive test result for the skin test or the IFN-γ ELISA can be challenging. Important information regarding how young animals can be tested and the meaning of a positive CMI test can only be answered by longitudinal studies. In this study 17 calves in year 1 and 13 calves in year 2 born to naturally MAP infected dams were skin tested every 90 days, blood collected for the IFN-γ ELISA and feces collected
for culture every 30 days. Tissues were collected for bacterial culture between 12 and 14 months of age.

During year 1, calves were separated from their dams at birth and maintained in a dry lot away from other cattle throughout the study. During year 2 calves were separated from birth and maintained in a dry lot until they were between 6 and 8 months of age then they were put out on pasture with other infected cattle. Preliminary results can be found in Appendix E. During year 1 the skin test identified 5 of the 10 infected calves by 6 months of age and 7 of 10 infected calves by 12 months of age. The skin test did not identify any culture negative calves by 6 months of age, but did identify 3 of 7 at 12 months of age. If a cut point of 0.2 was used, the IFN-γ ELISA identified all of the infected calves and 4 of the 7 culture negative calves at 12 months of age.

At year 2 the skin test identified all 13 calves as positive regardless of their culture status. Similar results were found with the IFN-γ ELISA results. More in depth analysis needs to be done, but it seems exposure of calves to infected cattle can affect the correlation of CMI test results to tissue culture status. Of course, this observation needs to be confirmed.

7.5 Recommendations for future research

Making PPD lots that have no batch to batch variation is extremely difficult. In order for a CMI diagnostic test to be useful, false positives must be minimized. A synthetic PPD with known quantities of antigens is an important area for future research.

Also, further understanding of what a positive CMI response means is needed. If infection must occur in order for an animal to have a positive CMI response, how often
do animals successfully overcome an infection? How long does an animal maintain a
detectable CMI response after the infection is gone?

For the gamma interferon ELISA, more work is needed to identify a better
positive control. Work also needs to be done looking at methods to add antigen out in the
field prior to transport to the laboratory.

Finally we need to know the herd sensitivity and specificity of these tests,
particularly if small subsets of animals are tested.
### APPENDIX A. SPECIFICITY DIFFERENCES IN ANTIGENS

Estimated specificities of the skin test and the IFN-γ ELISA using johnin, old tuberculin, 3 johnin PPD’s, an avian PPD, PBS and no injection in cattle and sheep.

<table>
<thead>
<tr>
<th></th>
<th>Cattle Skin test</th>
<th>Sheep Skin test</th>
<th>Cattle IFN-γ ELISA</th>
<th>Sheep IFN-γ ELISA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johnin OT</td>
<td>------</td>
<td>88.9 %±6.6%</td>
<td>------</td>
<td>88.4%±4.8%</td>
</tr>
<tr>
<td>Johnin 9801</td>
<td>97.6%±2.1%</td>
<td>98.9 %±1.0%</td>
<td>98.5%±1.3%</td>
<td>98.2%±1.9%</td>
</tr>
<tr>
<td>Johnin 0201</td>
<td>93.9%±3.9%</td>
<td>------</td>
<td>98.9%±1.6%</td>
<td>------</td>
</tr>
<tr>
<td>Johnin 0202</td>
<td>84.2%±3.8%</td>
<td>84.9 %±5.4%</td>
<td>95.3%±2.2%</td>
<td>94.8%±3.4%</td>
</tr>
<tr>
<td>M avium PPD</td>
<td>------</td>
<td>------</td>
<td>97.2%±1.7%</td>
<td>97.1%±2.5%</td>
</tr>
<tr>
<td>PBS only</td>
<td>98.7%±1.5%</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>(1/227)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No injection</td>
<td>(1/116)</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
</tbody>
</table>
APPENDIX B. REPEATABILITY OF THE SKIN TEST

The repeatability of the skin test in a Mycobacterium avium subsp. paratuberculosis-infected sheep flock tested 12 months apart.

<table>
<thead>
<tr>
<th>Skin test result</th>
<th>Year 1</th>
<th>Year 2</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>Negative</td>
<td>220</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>Positive</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>Positive</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>Negative</td>
<td>31*</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>344</td>
<td></td>
</tr>
</tbody>
</table>

*Two of the 31 sheep tested positive to the AGID test at year 2
APPENDIX C. FALSE POSITIVE IFN-γ ELISA RESULTS IN A SHEEP FLOCK

Optical density readings and test interpretation from the IFN-γ ELISA from a flock with severe oral lesions due to foxtail hay

<table>
<thead>
<tr>
<th>Tag #</th>
<th>Skin test</th>
<th>NS</th>
<th>NS</th>
<th>SEA</th>
<th>SEA</th>
<th>L3</th>
<th>L3</th>
<th>L4</th>
<th>L4</th>
<th>L3 result</th>
<th>L4 result</th>
</tr>
</thead>
<tbody>
<tr>
<td>B442</td>
<td>0</td>
<td>0.29</td>
<td>0.30</td>
<td>0.26</td>
<td>0.19</td>
<td>-0.15</td>
<td>-0.18</td>
<td>-0.17</td>
<td>-0.18</td>
<td>Invalid*</td>
<td>NA</td>
</tr>
<tr>
<td>Y48</td>
<td>7</td>
<td>0.55</td>
<td>0.54</td>
<td>2.59</td>
<td>2.53</td>
<td>2.13</td>
<td>2.13</td>
<td>1.18</td>
<td>1.15</td>
<td>Invalid</td>
<td>NA</td>
</tr>
<tr>
<td>Y14</td>
<td>6</td>
<td>0.29</td>
<td>0.31</td>
<td>0.28</td>
<td>0.32</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.02</td>
<td>Invalid</td>
<td>NA</td>
</tr>
<tr>
<td>Y38</td>
<td>0</td>
<td>2.56</td>
<td>2.33</td>
<td>0.73</td>
<td>0.62</td>
<td>0.65</td>
<td>0.54</td>
<td>-2.22</td>
<td>-2.22</td>
<td>Invalid</td>
<td>NA</td>
</tr>
<tr>
<td>Y28</td>
<td>0</td>
<td>0.78</td>
<td>0.74</td>
<td>1.14</td>
<td>1.07</td>
<td>0.23</td>
<td>0.21</td>
<td>-0.37</td>
<td>-0.40</td>
<td>Invalid</td>
<td>NA</td>
</tr>
<tr>
<td>Y46</td>
<td>0</td>
<td>0.21</td>
<td>0.20</td>
<td>0.62</td>
<td>0.64</td>
<td>-0.13</td>
<td>-0.13</td>
<td>-0.11</td>
<td>-0.12</td>
<td>Invalid</td>
<td>NA</td>
</tr>
<tr>
<td>Y18</td>
<td>0</td>
<td>0.45</td>
<td>0.44</td>
<td>0.63</td>
<td>0.57</td>
<td>0.06</td>
<td>0.03</td>
<td>-0.33</td>
<td>-0.34</td>
<td>Invalid</td>
<td>NA</td>
</tr>
<tr>
<td>B429</td>
<td>0</td>
<td>0.87</td>
<td>0.88</td>
<td>2.32</td>
<td>2.17</td>
<td>-0.08</td>
<td>-0.11</td>
<td>0.94</td>
<td>0.99</td>
<td>Invalid</td>
<td>NA</td>
</tr>
<tr>
<td>B428</td>
<td>0</td>
<td>0.55</td>
<td>0.54</td>
<td>0.57</td>
<td>0.54</td>
<td>-0.07</td>
<td>-0.06</td>
<td>0.56</td>
<td>0.48</td>
<td>Invalid</td>
<td>NA</td>
</tr>
<tr>
<td>B470</td>
<td>0</td>
<td>0.07</td>
<td>0.06</td>
<td>0.78</td>
<td>0.69</td>
<td>0.20</td>
<td>0.21</td>
<td>0.02</td>
<td>0.03</td>
<td>Posa</td>
<td>Neg</td>
</tr>
<tr>
<td>R407</td>
<td>4</td>
<td>0.13</td>
<td>0.11</td>
<td>0.87</td>
<td>0.88</td>
<td>0.57</td>
<td>0.50</td>
<td>0.52</td>
<td>0.50</td>
<td>Pos</td>
<td>Pos</td>
</tr>
<tr>
<td>B403</td>
<td>0</td>
<td>0.05</td>
<td>0.06</td>
<td>0.80</td>
<td>0.76</td>
<td>0.08</td>
<td>0.09</td>
<td>0.91</td>
<td>0.93</td>
<td>Neg</td>
<td>Pos</td>
</tr>
<tr>
<td>Y26</td>
<td>0</td>
<td>0.07</td>
<td>0.06</td>
<td>0.64</td>
<td>0.62</td>
<td>1.63</td>
<td>1.48</td>
<td>0.03</td>
<td>0.04</td>
<td>Pos</td>
<td>Neg</td>
</tr>
<tr>
<td>R411</td>
<td>0</td>
<td>0.05</td>
<td>0.04</td>
<td>0.89</td>
<td>0.89</td>
<td>0.07</td>
<td>0.06</td>
<td>0.98</td>
<td>1.03</td>
<td>Neg</td>
<td>Pos</td>
</tr>
<tr>
<td>B408</td>
<td>0</td>
<td>0.06</td>
<td>0.05</td>
<td>0.46</td>
<td>0.49</td>
<td>0.07</td>
<td>0.08</td>
<td>0.04</td>
<td>0.05</td>
<td>Neg</td>
<td>Neg</td>
</tr>
<tr>
<td>OR42</td>
<td>0</td>
<td>0.05</td>
<td>0.05</td>
<td>1.56</td>
<td>1.58</td>
<td>0.24</td>
<td>0.24</td>
<td>0.01</td>
<td>0.00</td>
<td>Pos</td>
<td>Neg</td>
</tr>
<tr>
<td>Y43</td>
<td>0</td>
<td>0.05</td>
<td>0.05</td>
<td>0.33</td>
<td>0.30</td>
<td>0.24</td>
<td>0.26</td>
<td>0.00</td>
<td>0.01</td>
<td>Pos</td>
<td>Neg</td>
</tr>
<tr>
<td>B404</td>
<td>0</td>
<td>0.05</td>
<td>0.05</td>
<td>1.24</td>
<td>1.42</td>
<td>1.45</td>
<td>1.60</td>
<td>0.01</td>
<td>0.01</td>
<td>Pos</td>
<td>Neg</td>
</tr>
<tr>
<td>Y39</td>
<td>0</td>
<td>0.05</td>
<td>0.05</td>
<td>2.20</td>
<td>2.16</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
<td>Neg</td>
<td>Neg</td>
</tr>
<tr>
<td>Y24</td>
<td>0</td>
<td>0.08</td>
<td>0.07</td>
<td>0.37</td>
<td>0.36</td>
<td>0.09</td>
<td>0.09</td>
<td>0.54</td>
<td>0.56</td>
<td>Neg</td>
<td>Pos</td>
</tr>
<tr>
<td>B464</td>
<td>0</td>
<td>0.11</td>
<td>0.10</td>
<td>0.29</td>
<td>0.28</td>
<td>0.01</td>
<td>0.01</td>
<td>-0.02</td>
<td>-0.01</td>
<td>Neg</td>
<td>Neg</td>
</tr>
<tr>
<td>B438</td>
<td>0</td>
<td>0.06</td>
<td>0.07</td>
<td>0.79</td>
<td>0.81</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>Neg</td>
<td>Neg</td>
</tr>
<tr>
<td>Y10</td>
<td>0</td>
<td>0.08</td>
<td>0.08</td>
<td>0.61</td>
<td>0.63</td>
<td>0.49</td>
<td>0.51</td>
<td>0.08</td>
<td>0.08</td>
<td>Pos</td>
<td>Neg</td>
</tr>
<tr>
<td>B430</td>
<td>0</td>
<td>0.05</td>
<td>0.05</td>
<td>0.98</td>
<td>0.91</td>
<td>0.17</td>
<td>0.16</td>
<td>0.73</td>
<td>0.68</td>
<td>Neg</td>
<td>Pos</td>
</tr>
<tr>
<td>B407</td>
<td>0</td>
<td>0.06</td>
<td>0.05</td>
<td>0.94</td>
<td>0.93</td>
<td>0.01</td>
<td>0.02</td>
<td>0.10</td>
<td>0.14</td>
<td>Neg</td>
<td>Neg</td>
</tr>
</tbody>
</table>

NS=Not stimulated background levels of IFN-γ  
SEA= Staphylococcus enterotoxin A  
L3= Johnin PPD 0202 (lot 3)  
L4 = Johnin PPD 0203 (lot 4)  
* Invalid means the background levels of IFN-γ are too high  
aOD readings at or above .2 are considered positive
APPENDIX D. ALTERNATIVE RESULTS USING THE TAGS PROGRAM TO ESTIMATE SENSITIVITY AND SPECIFICITY WITHOUT A GOLD STANDARD

Estimated sensitivity and specificity of the AGID, ELISA and skin test using the maximum likelihood program TAGS. These results were calculated using the test results from the sheep tested in Chapter 3.

<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity</th>
<th>95% CI</th>
<th>Specificity</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGID</td>
<td>0.0837</td>
<td>0.0647 - 0.1076</td>
<td>1</td>
<td>Not determined</td>
</tr>
<tr>
<td>ELISA</td>
<td>0.0830</td>
<td>0.0642 - 0.1068</td>
<td>0.9973</td>
<td>0.9898 - 0.9993</td>
</tr>
<tr>
<td>Skin Test</td>
<td>0.6954</td>
<td>0.6073 - 0.7712</td>
<td>0.9879</td>
<td>0.9711 - 0.9950</td>
</tr>
</tbody>
</table>

Correlations among the diagnostic tests. The correlation between the AGID and the ELISA caused this model to be suspect.

<table>
<thead>
<tr>
<th></th>
<th>AGID Vs</th>
<th>AGID Vs.</th>
<th>ELISA vs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELISA</td>
<td>Correlation</td>
<td>-0.0456</td>
<td>-0.0262</td>
</tr>
<tr>
<td>Skin test</td>
<td></td>
<td>Correlation</td>
<td>0.2311</td>
</tr>
</tbody>
</table>
APPENDIX E. PRELIMINARY RESULTS OF THE SKIN TEST AND IFN-γ ELISA FROM CALVES BORN TO MAP INFECTED DAMS

Year 1 skin test and IFN-γ ELISA results from calves born to naturally MAP infected dams.

<table>
<thead>
<tr>
<th>Calf #</th>
<th>IFN-γ (OD)</th>
<th>IFN-γ (OD)</th>
<th>IFN-γ (OD)</th>
<th>Skin Test (mm)</th>
<th>Skin Test (mm)</th>
<th>Skin Test (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4-6 mo</td>
<td>7-9 mo</td>
<td>12-13 mo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.365</td>
<td>0.290</td>
<td>1.046</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>0.883</td>
<td>-</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>0.644</td>
<td>0.775</td>
<td>1.309</td>
<td>3.5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>21</td>
<td>-</td>
<td>-</td>
<td>2.948</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>-</td>
<td>0.165</td>
<td>0.552</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>52</td>
<td>-</td>
<td>-</td>
<td>0.691</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>53</td>
<td>0.191</td>
<td>-</td>
<td>2.631</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>56</td>
<td>-</td>
<td>2</td>
<td>0.459</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>62</td>
<td>0.308</td>
<td>0.441</td>
<td>0.880</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>57</td>
<td>-</td>
<td>0.375</td>
<td>1.789</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Ave</td>
<td>0.377</td>
<td>0.494</td>
<td>1.32</td>
<td>2.7</td>
<td>2.75</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calf #</th>
<th>IFN-γ (OD)</th>
<th>IFN-γ (OD)</th>
<th>IFN-γ (OD)</th>
<th>Skin Test (mm)</th>
<th>Skin Test (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4-6 mo</td>
<td>7-9 mo</td>
<td>12-13 mo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.279</td>
<td>0.119</td>
<td>0.354</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
<td>0.567</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>28</td>
<td>0.212</td>
<td>0.154</td>
<td>0.171</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>34</td>
<td>-</td>
<td>-</td>
<td>0.131</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>71</td>
<td>-</td>
<td>0.124</td>
<td>0.313</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>235</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td>-</td>
<td>0.233</td>
<td>0.508</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ave</td>
<td>0.246</td>
<td>0.158</td>
<td>0.338</td>
<td>2.33</td>
<td></td>
</tr>
</tbody>
</table>
Year 2 skin test and IFN-γ ELISA results from calves born to naturally MAP infected dams. After the calves were 6 months old they were moved to a pasture with MAP infected cows.

<table>
<thead>
<tr>
<th>Calf No</th>
<th>Culture Results</th>
<th>&lt;4 mo</th>
<th>4-6 mo</th>
<th>6-8 mo</th>
<th>10-12 mo</th>
<th>12+</th>
<th>&lt;4 mo</th>
<th>4-6 mo</th>
<th>6-8 mo</th>
<th>10-12 mo</th>
<th>12+</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>31</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>0</td>
<td>++</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>34</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>35</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>38</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>42</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>21</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>25</td>
<td>++</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>++</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>37</td>
<td>++</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>39</td>
<td>++</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>43</td>
<td>++</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>44</td>
<td>++</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>0</td>
<td>0</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

Science of course is no longer an individual endeavor and many people have been very generous with their time and knowledge. There is simply no room to thank everyone that has touched my career here at Iowa State.

I do want to recognize the researchers in Johne’s disease who have come before me. They have left a foot print and a guide book with their papers that has given me a great platform to build on.

I want to thank the producers who have taken the time out of their busy schedules so that I could test their flocks and herds.

I would like to thank Ron Griffith, my major professor who has simply been just wonderful. Someday I hope I can respond in kind.

I would like to thank Judy Stabel, who has taught be so much and has become a very close friend as well.

I would like to thank John Bannantine, Rich Evans, and Annette O’Connor who agreed to be on my committee and who have offered me guidance.

And finally to Tony, my husband who worried and stressed about my well being more than I did. Nobody could have a better partner than you.