# STATISTICAL MODELING OF LONGITUDINAL DATA FROM DAIRY HERDS: *MYCOBACTERIUM AVIUM* SUBSP. *PARATUBERCULOSIS*

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Rebecca Lee Smith

August 2011

© 2011 Rebecca Lee Smith

# STATISTICAL MODELING OF LONGITUDINAL DATA FROM DAIRY HERDS: *MYCOBACTERIUM AVIUM* SUBSP. *PARATUBERCULOSIS*

Rebecca Lee Smith, D.V.M., M.S., Ph. D.

Cornell University 2011

*Mycobacterium avium* subsp. *paratuberculosis* (MAP) causes infections in ruminants characterized by long latent periods, imperfect diagnostic tests, and production effects sufficient to result in early culling. In order to optimize control of this pathogen, the exact effects of both the infection and the control strategies must be understood. The goal of this dissertation is to statistically analyze various aspects of MAP in dairy herds, including the production effects of infection, its distribution in the dairy herd environment, and the transmission rates of animals shedding MAP. For this purpose, longitudinal data sets from commercial dairy herds have been analyzed, allowing for a more thorough understanding of MAP in typical farms. Cows in the high-shedding category (>30 cfu/g of MAP in feces) were found to produce approximately 4 kg less milk per day, on average, and to have higher culling rates and lower calving rates than non-shedding cows. In addition, the number of high-shedding animals in a pen was positively correlated with the amount of MAP cultured from the environment in that pen. In contrast, low-shedding cows were found to have higher culling rates than non-shedding cows, but no significant difference in calving rates or milk production. The average amount of fecal shedding in the herd was found to be predictive of both the odds of MAP being found in environmental samples and the amount of MAP in those samples, but environmental sampling was not found to be a sensitive herd-level diagnostic test. These results will enable optimization of economic models for MAP

control by providing quantitative estimates of the effect of MAP on commercial dairy farms. In addition, it was found that reversible-jump Markov Chain Monte Carlo models are unable to estimate transmission rates for MAP using current longitudinal data sets, due to the large amount of missing data.

## BIOGRAPHICAL SKETCH

Rebecca Smith is a D.V.M. finishing a Ph. D. in the Department of Population Medicine and Diagnostic Sciences at the Cornell University College of Veterinary Medicine, where she is supervised by Drs. Yrjo Gröhn and Ynte Schukken, with additional supervision by Dr. Loren Tauer from the Charles H. Dyson School of Applied Economics and Management and by Dr. Robert Strawderman from the Department of Biological Statistics and Computational Biology. She received a B.A. in biology, with a minor in English, from Gustavus Adolphus College in 2001, graduating *cum laude* and as a member of the Guild of St. Ansgar. In 2005, she received her D.V.M. from Cornell University College of Veterinary Medicine, where she was recognized for her senior seminar project entitled "Environmental sampling to determine herd infection status with *Mycobacterium avium* subsp. *paratuberculosis*". Kansas State University awarded her an M.S. in biosecurity and risk analysis in 2007 for the thesis "Biosecurity and risk analysis for cow-calf enterprises: a simulation model for Bovine Viral Diarrhea Virus". The Kansas State University chapter of Phi Zeta recognized her with a Mahlon Vorhies Production Animal Award for her project "A stochastic risk analysis model for the spread of bovine viral diarrhea virus in cow-calf herds" in 2007. As a graduate student at Cornell, she has received a Travel Award from the Johne's Disease Integrated Program, and has been supported by Cornell University's Institutional Training Grant (T32 RR007059). She has served as an invited reviewer for the *Journal of Dairy Science* and is a member of the American Veterinary Medical Association. She was recently awarded a K01 Career Development Grant under the National Council on Research Resources (NIH) for "Mathematical and statistical modeling of mycobacterial infections", a continuation of her thesis research.

Dedicated to my family and friends, for their patience,

and to Alex, for his timing

ACKNOWLEDGMENTS

The USDA (Cooperative State Research, Education and Extension Service, Washington, D.C.) Award Number 2008-35204-04627 provided funding for these studies, as did the USDA-Agricultural Research Service (Agreements. 58-1265-3-155, 58-1265-3-156, 58-1265-3-158, and 58-1265-4-020) for the Regional Dairy Quality Management Alliance (RDQMA) and the Johne's Disease Integrated Program (JDIP, USDA contract 45105).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

| | |
|---|---|
| AR | acceptance ratio |
| cfu/g | colony forming units/gram |
| CI | confidence interval |
| DHIA | Dairy Herd Improvement Association |
| DIM | days in milk |
| DNB | do not breed |
| EC | environmental culture |
| ELISA | enzyme-linked immunosorbent assay |
| FC | fecal culture |
| HEY | Herrold's egg yolk media |
| IPCW | inverse probability of censoring weighted |
| JD | Johne's disease |
| JDDHP | Johne's Disease Demonstration Herd Project |
| JDIP | Johne's Disease Integrated Program |
| kg | kilogram |
| LS | linear score |
| MAP | *Mycobacterium avium* subsp. *paratuberculosis* |
| MCMC | Markov Chain Monte Carlo |
| ME | mature equivalent |
| NAHMS | National Animal Health Monitoring Service |
| NES | NAHMS environmental set |
| NPV | net present value |
| NY | New York |
| RDQMA | Regional Dairy Quality Management Alliance |
| rjMCMC | reversible-jump Markov Chain Monte Carlo |

| | |
|---|---|
| PA | Pennsylvania |
| PI | prediction interval |
| VT | Vermont |

# INTRODUCTION

## STATISTICAL MODELING OF LONGITUDINAL DATA FROM DAIRY HERDS: *MYCOBACTERIUM AVIUM* SUBSP. *PARATUBERCULOSIS*

The US dairy industry has been faced in recent years with a difficult economic situation. The low price of milk and the high price of fuel and feed have decreased the profit margins for dairy production, sometimes to the point of net losses (Karszes et al., 2009; Knoblauch et al., 2009). In these circumstances especially, disease control recommendations for dairy producers should be economically justified.

Among the major endemic pathogens the US dairy industry would like to control is *Mycobacterium avium* subsp. *paratuberculosis* (MAP) (2007), a slow-growing, gram-negative species of the common mycobacterial genus. Cattle and other ruminants infected with MAP generally experience a long latent period, with no clinical or subclinical signs, followed by gradual, chronic progression to Johne's disease (JD) (Whitlock et al., 2000). Subclinical JD is known to cause decreased milk production (Nielsen et al., 2008; Smith et al., 2009), and clinical JD results in wasting (Kennedy and Benedictus, 2001), which decreases slaughter value. Between milk loss and low slaughter value, JD is thought to cost the US dairy industry $200 million/year (Ott et al., 1999).

Control of MAP, however, is difficult, and eradication may be impossible. In the dairy environment, MAP can survive up to 6 months, especially in areas of fecal concentration like manure slurry (Whittington et al., 2004; Whittington et al., 2005). Animals latently infected with MAP are not detectable antemortem; animals with subclinical JD may be detected with a variety of diagnostic tests (Collins et al., 2006),

but the diagnostic sensitivity of these tests is low (Whitlock et al., 2000). All of these factors contribute to the persistence of MAP in dairy farms, even with the toughest (and most expensive) control programs (Raizman et al., 2006; Wells et al., 2008). In fact, the cost/benefit ratio of strict control may be too high to be feasible; the economically optimal MAP control strategy in dairy herds may very well be less stringent than the biologically optimal control strategy (Lu et al., 2010). In order to determine the economically optimal MAP control strategy, however, the exact effect of MAP infection on production and the exact effect of MAP control on infection must both be estimated.

As JD is chronic, slow to develop, and difficult to identify, the best estimates will come from repeated testing and longitudinal production data. Statistical analysis of longitudinal data requires distinguishing covariates as time-dependent (such as disease status), time-independent (such as breed), or random effects (such as herd); response variables may be repeated measures (such as milk production or test results) or event data (such as infection or calving). When all variables are known or can be reasonably estimated, classical statistical approaches can be used to estimate parameters; for example, mixed models of repeated measures can be handled with an autocorrelation structure (Greene, 2008). When crucial variables are unknown, however, the classical statistical approach breaks down. Bayesian methods, however, are capable of handling unknown variables, however, by treating them as nuisance parameters (Carlin and Louis, 2009; Gelman et al., 2004). In the following work, classical statistics have been used wherever they are appropriate (Chapters 1-3); in

Chapters 4 and 5, classical methods were insufficient, and a Bayesian method (reversible-jump Markov Chain Monte Carlo modeling) was used.

The objective of the research described in this dissertation is to provide accurate estimates of the effect of MAP on dairy herd production, specifically milk production (Chapter 1) and reproduction and culling (Chapter 2). This research also attempts to estimate the efficacy of MAP control programs, specifically, environmental testing (Chapter 3) and vaccination (Chapter 5). In the process, a generalizable model is presented that can estimate the transmission rate of MAP in a given dairy herd (Chapter 4), which will allow estimation of the efficacy of MAP control programs focusing on decreasing transmission rates, such as those focused on hygiene. The results of this research can be directly applied within economic models for MAP, allowing for more accurate predictions of the economically optimal MAP control strategy.

## REFERENCES

2007. NAHMS 2007: Johne's Disease on U.S. Dairy Operations. Pages 1-4 USDA:APHIS:VS:NAHMS, Fort Collins, CO.

Carlin, B. P., and T. A. Louis. 2009. Bayesian Methods for Data Analysis. 3rd ed. Taylor & Francis Group, LLC, Boca Raton, FL.

Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, Jr., and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229(12):1912-1919.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2004. Bayesian Data Analysis. 2nd ed. Chapman & Hall/CRC, Boca Raton, FL.

Greene, W. H. 2008. Models for Panel Data. Pages 180-251 in Econometric analysis. D. Alexander, ed. Pearson Education Inc., Upper Saddle River, NJ.

Karszes, J., W. A. Knoblauch, and L. D. Putnam. 2009. Dairy Business Summary: New York large herd farms, 300 cows or larger 2008. Cornell University.

Kennedy, D., and G. Benedictus. 2001. Control of Mycobacterium avium subsp. paratuberculosis infection in agricultural species. Rev. Sci. Technol. 20:151-179.

Knoblauch, W. A., L. D. Putnam, M. Kiraly, and J. Karszes. 2009. Dairy Farm Business Summary: New York small herd farms, 80 cows or fewer 2008. Cornell University.

Lu, Z., Y. H. Schukken, R. L. Smith, and Y. T. Grohn. 2010. Stochastic simulations of a multi-group compartmental model for Johne's disease on US dairy herds with test-based culling intervention. J. Theo. Bio. 264(4):1190-1201.

Nielsen, S. S., M. A. Krogh, and C. Enevoldsen. 2008. Time to the occurence of a decline in milk production in cows with various paratuberculosis antibody profiles. J. Dairy Sci. 92:149-155.

Ott, S. L., S. J. Wells, and B. A. Wagner. 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. Prev. Vet. Med. 40(3-4):179-192.

Raizman, E. A., S. J. Wells, S. M. Godden, J. P. Fetrow, K. Friendshuh, and O. J. Michael. 2006. Characterization of Minnesota dairy herds participating in a Johne's disease control program and evaluation of the program risk assessment tool. Prev. Vet. Med. 75:22-33.

Smith, R. L., Y. T. Grohn, A. K. Pradhan, R. H. Whitlock, J. S. Van Kessel, J. M. Smith, D. R. Wolfgang, and Y. H. Schukken. 2009. A longitudinal study for the impact of Johne's Disease status on milk production in individual cows. J. Dairy Sci. 92(6):2653-2661.

Wells, S. J., W. L. Hartmann, and P. L. Anderson. 2008. Evaluation of progress made by dairy and beef herds enrolled in the Minnesota Johne's Disease Control Program. J. Am. Vet. Med. Assoc. 233(12):1920-1926.

Whitlock, R. H., S. J. Wells, R. W. Sweeney, and T. J. Van. 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): sensitivity and specificity of each method. Vet. Microbiol. 77(3-4):387-398.

Whittington, R. J., I. B. Marsh, and L. A. Reddacliff. 2005. Survival of Mycobacterium avium subsp. paratuberculosis in dam water and sediment. Appl. Environ. Microbiol. 71(9):5304-5308.

Whittington, R. J., D. J. Marshall, P. J. Nicholls, I. B. Marsh, and L. A. Reddacliff. 2004. Survival and dormancy of Mycobacterium avium subsp. paratuberculosis in the environment. Appl. Environ. Microbiol. 70(5):2989-3004.

CHAPTER 1

# A LONGITUDINAL STUDY FOR THE IMPACT OF JOHNE'S DISEASE STATUS ON MILK PRODUCTION IN INDIVIDUAL COWS

*originally published as:*

*The impact of Johne's Disease status on milk production in individual cows. Smith, R.L.; Grohn, Y.T.; Pradhan, A.K.; Whitlock, R.H.; Van Kessel, J.S.; Smith, J.M.; Wolfgang, D.R.; Schukken, Y.H. J. Dairy Sci. 2009. 92(6): 2653-2661.*

***Abstract***

Longitudinal data from three commercial dairy herds in the Northeast United States were collected from 2004 to 2007. Johne's Disease status, as indicated by *Mycobacterium avium* subsp. *paratuberculosis* infection levels, was determined through quarterly ELISA serum testing, biannual fecal culture, and culture of tissues at slaughter. Milk production data were collected from the Dairy Herd Improvement Association. The effect of Johne's Disease status on milk production was analyzed using a mixed linear model with an autocorrelation random effect structure. Infected animals produced more milk than uninfected cows before they began shedding *Mycobacterium avium* subsp. *paratuberculosis*. Cows infected with *Mycobacterium avium* subsp. *paratuberculosis* had monthly decreases of 0.05 to 1 kg in daily milk production relative to uninfected animals, with greater decreases in progressive disease categories. Animals with fecal culture results of more than 30 cfu/gram produced approximately 4 kg less milk per day compared to uninfected cows. These results will be valuable in calculating the economic effect of Johne's Disease.

***Introduction***

Johne's Disease (JD) is a chronic disease of ruminants caused by intestinal infection with the pathogen *Mycobacterium avium* subsp. *paratuberculosis* (MAP). The pathogen is pervasive on US dairy farms, with a herd prevalence of approximately 68% (NAHMS, 2007). Infection with MAP typically occurs in calves (Clarke, 1997), which then enter a latent, non-shedding stage of varying length. This latent stage is followed by a period of low and intermittent shedding of MAP with no obvious clinical symptoms (Whitlock et al., 2000). If allowed to progress, clinically apparent JD may develop, with a high level of MAP shedding (Whitlock et al., 2000).

The most commonly used ante-mortem tests for identifying MAP shedding are fecal culture and serum ELISA (NAHMS, 2007). Culture techniques are used to determine both the presence and the magnitude of MAP shed in the feces, while serum ELISA testing is used to detect an immune response to infection. Both of these tests are sensitive in identifying high-shedding animals, but are much less sensitive for low-shedding animals and often fail to detect latent infections (Eamens et al., 2000; Whitlock et al., 2000). Identification of infected animals is best achieved by repeated sampling (van Schaik et al., 2003). Post-mortem culture of tissues, including intestinal epithelium and lymph nodes, can identify infection at any stage, including latency, but the procedure is more sensitive in animals with advanced infections (Huda and Jensen, 2003).

Structured control programs often include immediate culling of high shedding animals to limit environmental contamination and the transmission of MAP to herdmates, particularly calves. However, a combination of imperfect diagnostic

techniques and slow development of clinical symptoms often results in delayed culling or retention of low-shedding animals. Additionally, low-shedding animals are frequently retained due to the high cost of purchasing replacements (Dorshorst et al., 2006).

Johne's Disease has been estimated to cost the US dairy industry $200 to $250 million annually (Ott et al., 1999). The cost of JD is manifested in a variety of ways, but milk production losses are the most insidious. Decreased milk production is believed to occur in sub-clinical and clinical animals (Kennedy and Benedictus, 2001), but the precise magnitude of the reduction has not been previously calculated. One study found no significant decrease in 305-day mature equivalent (**ME**) milk production for a small cohort of animals positive for JD by ELISA and/or fecal culture (Johnson et al., 2001). However, in that study, animals were followed for only 18 months with tests at the onset and finish of the study. Given the low sensitivity of the diagnostics to low-shedding animals, there was a high probability of misclassification of control animals. Another study noticed a significant decrease in 305-day ME in fecal culture-positive animals (Wilson et al., 1993). Other studies found that total milk production was decreased over the course of a lactation in fecal culture-positive animals (Raizman et al., 2007), and that ELISA-positive animals had lower ME 305-day production and lifetime milk production (Lombard et al., 2005) compared with test-negative animals. The magnitude of continuous test results from the milk ELISA for MAP antibodies have also been found to be significantly related to the shape of the lactation curves (Kudahl et al., 2004).

Each of these 3 studies, however, based their results entirely on a single test per animal and were thus potentially biased by the high probability of false-negative results. False-negative results would misclassify an infected animal as a control and thereby reduce the calculated impact of infection on the outcome of interest. Also, an infection status change, for example, from low to high shedding, during the data collection period would not have been taken into account in these studies. Now, more precise estimates of production life can be estimated using test day milk yield models, which account for random cow effects, repeated observation, and time-dependent covariates indicating disease status (Wilson et al., 2004). Because of the high uncertainty associated with the previous estimates, the reduction of milk production attributed to MAP infection contributes most of the uncertainty in economic models for the cost of JD (Losinger, 2005).

Without accurate milk loss predictors, the economic impact of MAP infection cannot be known with certainty. The objective of this study is therefore to estimate the effect of JD status on individual cow milk production using longitudinal data collected over a 4-year period from 3 US dairy herds.

***Materials and Methods***

In 2004, the Regional Dairy Quality Management Alliance (RDQMA) identified 1 commercial dairy herd in 3 of its member states (NY, VT, and PA) to serve as a longitudinal study herd. The details of this study have been described previously (Pradhan et al., 2009). Briefly, these 3 herds, which are endemically MAP-infected and consist primarily of Holstein cattle, were visited on a quarterly basis to collect samples, and all monthly production, breeding, and health records were

obtained through the Dairy Herd Improvement Association (DHIA). Serum samples were collected quarterly and fecal samples were collected biannually from each adult animal in each herd, and all samples were shipped overnight to The University of Pennsylvania Johne's Laboratory. In addition, samples (including serum, feces, intestinal epithelium, and intestinal lymph nodes) were collected from almost all animals sent to slaughter and a majority of animals that died on-farm during the study period. These samples were refrigerated at the slaughterhouse and shipped within one day of collection. Serum samples were tested by the ParaChek® by Prionics (Prionics USA Inc., La Vista, NE; formerly CSL/Biocor) enzyme-linked immunosorbent assay (ELISA) for antibody reactions to MAP antigens. Fecal and tissue samples were tested by four-tube fecal culture for presence of viable MAP organisms (Pradhan et al., 2009). Diagnostic results were reported as positive or negative for ELISA and as cfu/g for culture.

Johne's Disease status (stage of MAP infection) was divided into 4 categories: uninfected, latent, low-shedding, and high-shedding (defined below). Uninfected animals were defined as animals for which all lifetime diagnostic tests were negative and for which there were at least 2 diagnostic test results reported. Animals with only 1 diagnostic test, if the result was negative, were removed from the analysis due to the lack of diagnostic sensitivity. Animals were assumed infected if at least 1 diagnostic test had a positive result. Johne's Disease status was assigned separately for each combination of animal and test day. All milk test days (days on which milk weights were recorded) prior to the first positive diagnosis were categorized as latent because the animals were assumed to be infected in calfhood. Animals with positive tissue

culture results were categorized as latent for all milk test days if no antemortem tests were positive. A cut-off value of 30 cfu/g was used to separate low- and high-shedding culture results, as representing the midpoint of the moderate-shedding level previously defined by Whitlock et al. (Whitlock et al., 2000). Milk test days after a fecal culture result of at least 30 cfu/g were categorized as high-shedding. All other milk test days in positive animals (falling after a positive diagnosis by ELISA or a fecal culture of less than 30 cfu/g) were categorized as low-shedding. Positive results on ELISA tests are generally considered to follow positive results on fecal culture (van Schaik et al., 2003), so ELISA-positive animals may be categorized as low-shedding rather than latent.

Parity was divided into first and later lactations to account for the different shape of the lactation curve in first parity animals. Calving season was categorized as winter (January to March), spring (April to June), summer (July to September), and fall (October to December). A separate variable was created for milk test day number (TDnr), indicating the number of milk test days for a given animal since the beginning of the study. It was assumed that 2 milk test days passed between lactations to account for time spent in the dry period, for which the US standard is approximately 60 days. In addition, a variable for the number of milk test days in the current JD status category (MthJ) was created for test-positive animals; MthJ was set equal to 0 for uninfected animals. Daily milk production (in kilograms), somatic cell linear score (**LS**, indicating the log of the somatic cell count), and days in milk in the current lactation (DIM) were used as reported by DHIA, with any recorded non-numeric and zero values for milk production and LS removed from the analysis.

Mixed model analysis was performed using the Mixed procedure in SAS® version 9 (SAS Institute Inc., Cary, NC, USA). Wilmink's correction (Wilmink, 1987) was used to model the lactation curve, with an interaction of parity and lactation curve to capture the different shape of the first-lactation curve. Parity, LS, herd, and calving season were included as fixed effects. Cow was included as a random variable, with first-order autoregression for each combination of cow and lactation using TDnr to identify the time lag between observations. The model structure was:

**Equation 1.1**

$$\begin{aligned} milk_{i,l,t} = {} & \beta_0 + \beta_{1,p}\, parity + \beta_2 DIM_{i,l,t} + \beta_3 \exp\left[-0.1 * DIM_{i,l,t}\right] + \beta_4 LS_{i,l,t} \\ & + \delta_{1,p} DIM_{i,l,t} * parity + \delta_{2,p} \exp\left[-0.1 * DIM_{i,l,t}\right] * parity + \beta_{5,m} season \\ & + \beta_{6,n} JD_{i,l,t} + \left(\rho_i \varepsilon_{i,t-1} + u_{i,l,t,p,m,n}\right) \end{aligned}$$

where the outcome is daily milk production (in kg), *i* indicates cow, *l* indicates the present lactation, and *t* indicates milk test day. $\beta_{1,p}$ is the fixed effect of the dichotomized $p^{th}$ parity ($p$ = 1, >1), $\beta_2$ is the effect of DIM, $\beta_3$ is the effect of Wilmink's correction, $\beta_4$ is the effect of LS, $\beta_{5,m}$ is the fixed effect of the $m^{th}$ season ($m$ = winter, spring, summer, fall), and $\beta_{6,n}$ is the fixed effect of the $n^{th}$ JD status ($n$ = uninfected, latent, low-shedding, high-shedding). The interaction coefficient $\delta_{1,p}$ is the effect of the interaction between DIM and the $p^{th}$ parity, while the interaction coefficient $\delta_{2,p}$ is the effect of the interaction between Wilmink's correction and the $p^{th}$ parity. The term $\rho_i \varepsilon_{i,t-1}$ provides the first-order autoregression between milk test days in individual cows, and $\mu_{i,l,t,p,m,n}$ is the error term for each test date. Differences in milk production were tested between each level of JD status using a standard F-test.

A subsequent model was built to take into account the monthly decrease in milk production within each JD status category, with the structure:

**Equation 1.2**

$$milk_{i,l,t} = \beta_0 + \beta_{1,p}\, parity + \beta_2 DIM_{i,l,t} + \beta_3 \exp\left[-0.1 * DIM_{i,l,t}\right] + \beta_4 LS_{i,l,t} + \delta_{1,p} DIM_{i,l,t} * parity + \delta_{2,p} \exp\left[-0.1 * DIM_{i,l,t}\right] * parity + \beta_{5,m} season + \beta_{6,n} JD_{i,l,t} + \beta_7 MthJ_{i,l,t} + \beta_{8,n} JD_{i,l,t} * MthJ_{i,l,t} + \left(\rho_i \varepsilon_{i,t-1} + u_{i,l,t,p,m,n}\right)$$

Where most parameters are as defined in equation 1.1 and $\beta_7$ is the linear effect of *MthJ*, $\beta_{8,n}$ is the fixed effect of the interaction between *MthJ* and the $n^{th}$ JD status, and *MthJ* is a measure of time spent at the current JD status (described above). This model is hereafter defined as the 'time-based' model. All effects were considered significant at the $\alpha = 0.05$ level.

***Results***

A description of the data available for analysis is included in Table 1.1. There were 24,474 observations available for analysis, representing 1,332 individual cows and 2,713 individual lactations (934 first lactations) for the period of January 2003 to December 2007. No fecal culture results had been reported for 2007 at the time of analysis. A total of 2,775 observations, approximately 10% of all observations, were deleted for lacking sufficient test results (at least one positive or two negative tests), or appropriate milk weights or LS (numeric, non-zero values).

Convergence criteria were met for all models. The results of the 2 models used are presented in Table 1.2. Classes not included in the variable list were considered baseline values and are included in the intercept. Tests of significance between JD

Table 1.1: Longitudinal data from three commercial dairy herds in the Northeast United States to assess the impact of Johne's Disease on milk production

| Herd | State | Years followed | Number of adult cows in study | Johne's Disease status | Average days in milk (standard deviation) | Average daily milk production in kg (standard deviation) |
|---|---|---|---|---|---|---|
| A | NY | 4.9 | 865 | Negative | 204 (141) | 38.07 (10.72) |
| | | | | Latent | 204 (139) | 41.23 (12.35) |
| | | | | Low-shedding | 262 (178) | 34.62 (11.04) |
| | | | | High-shedding | 234 (123) | 30.42 (13.39) |
| B | PA | 3.3 | 226 | Negative | 200 (136) | 32.67 (10.56) |
| | | | | Latent | 190 (127) | 32.33 (7.7) |
| | | | | Low-shedding | 260 (217) | 31.53 (10.42) |
| | | | | High-shedding | 253 (153) | 26.42 (8.3) |
| C | VT | 2.5 | 241 | Negative | 206 (135) | 27.61 (10.32) |
| | | | | Latent | 171 (106) | 30.54 (11.73) |
| | | | | Low-shedding | 224 (138) | 27.51 (11.59) |
| | | | | High-shedding | 185 (107) | 23.86 (10) |
| Total | | | 1,332 | | 205 (141) | 35.77 (11.5) |

| Herd | Number of cows in category | Average parity (standard deviation) | Total number of milk test results | Number of first-lactation results |
|---|---|---|---|---|
| A | 802 | 2.1 (1.2) | 15,317 | 6,096 |
| | 62 | 2.8 (1.5) | 1,211 | 222 |
| | 49 | 3.3 (1.5) | 596 | 70 |
| | 8 | 3.2 (1) | 45 | 0 |
| B | 214 | 2.3 (1.3) | 3,623 | 1,441 |
| | 11 | 1.9 (0.8) | 62 | 20 |
| | 6 | 3.2 (1.2) | 60 | 5 |
| | 2 | 4.1 (1.4) | 25 | 0 |
| C | 210 | 2.1 (1.2) | 3,072 | 1,228 |
| | 24 | 2.3 (1.1) | 142 | 43 |
| | 29 | 2.4 (1) | 304 | 56 |
| | 3 | 3.3 (0.6) | 17 | 0 |
| | N/A | 2.2 (1.3) | 24,474 | 9,181 |

Table 1.2: Results of 2 linear mixed models for daily milk production (in kg). In the first model, only the categorical Johne's Disease (JD) status of the animal at the time of the milk test is considered. In the second model, the number of months spent in the JD status category is added to the model (JD time=0 for uninfected animals). Herd was included as a fixed effect in the model, but is not reported here.

| Variable | | Categorical | | Time-Based | |
|---|---|---|---|---|---|
| | *Class* | Parameter | t (P-value) | Parameter | t (P-value) |
| Intercept | | 33.38 | 70.79 (<0.001) | 33.29 | 69.70 (<0.001) |
| Parity (base is first lactation) | *> 1 lactation* | 11.15 | 29.08 (<0.001) | 11.23 | 29.09 (<0.001) |
| Days In Milk (DIM) | | -0.03 | -25.06 (<0.001) | -0.03 | -25.00 (<0.001) |
| $e^{-0.1*DIM}$ | | -24.45 | -28.28 (<0.001) | -24.45 | -28.30 (<0.001) |
| DIM*parity (base is first lactation) | *> 1 lactation* | -0.04 | -26.64 (<0.001) | -0.04 | -26.65 (<0.001) |
| $e^{-0.1*DIM}$*parity (base is first lactation) | *> 1 lactation* | -4.30 | -4.01 (<0.001) | -4.48 | -4.17 (<0.001) |
| Linear Score | | -0.79 | -29.02 (<0.001) | -0.78 | -28.70 (<0.001) |
| Calving season (base is winter) | *Fall* | -0.48 | -1.49 (0.181) | -0.50 | -1.53 (0.170) |
| | *Spring* | -0.43 | -1.26 (0.247) | -0.38 | -1.08 (0.314) |
| | *Summer* | -1.44 | -4.32 (0.004) | -1.41 | -4.19 (0.004) |
| JD status (base is uninfected) | *latent* | 2.30 | 5.06 (<0.001) | 3.15 | 4.73 (<0.001) |
| | *Low-shedding* | 0.20 | 0.37 (0.709) | 0.86 | 1.20 (0.234) |
| | *High-shedding* | -3.70 | -1.94 (0.056) | 0.84 | 0.33 (0.744) |
| JD time | | | | -1.12 | -2.54 (0.011) |
| JD time*JD status (base is high-shedding) | *latent* | | | 1.08 | 2.44 (0.015) |
| | *Low-shedding* | | | 1.05 | 2.36 (0.018) |

status levels, based on each of the models, are shown in Table 1.3. All meaningful comparisons are included. Inclusion of a linear variable for the cfu/gm recorded in the most recent positive fecal culture resulted in a small, non-significant parameter estimate, and was therefore removed from the model.

Table 1.3: Comparison tests for the effect of JD status levels on daily milk production. The categorical model only considers the JD status at the time of the milk test. The time-based model also considers the effect of the number of months the animal has spent in the JD status category (*MthJ*)

| | Categorical | | | Time-Based | | |
|---|---|---|---|---|---|---|
| Comparison | Difference (kg) | F | p-value | Difference (kg) | F | p-value |
| Latent vs. negative | +2.3 | 5.06 | <0.001 | +3.15 -0.04**MthJ* | 4.73 | <0.001 |
| Latent vs. low-shedding | +2.1 | 13.40 | <0.001 | +2.29 +0.03**MthJ* | 0.15 | 0.695 |
| Latent vs. high-shedding | +6.0 | 9.82 | 0.003 | +2.31 +2.2**MthJ* | 5.94 | 0.015 |
| Low-shedding vs. negative | +0.20 | 0.37 | 0.709 | +0.86 -0.07**MthJ* | 1.20 | 0.234 |
| Low-shedding vs. high-shedding | +3.9 | 3.99 | 0.050 | +0.02 +2.17**MthJ* | 5.59 | 0.018 |
| High-shedding vs. negative | -3.7 | -1.94 | 0.056 | +0.84 -1.12**MthJ* | -2.54 | 0.011 |

Several cows had recorded lactation lengths that were much longer than typical; although the expected lactation length is approximately one year, 2,828 observations were more than 365 DIM and 104 observations were more than 730 DIM. This is likely due to a failure to report late-term abortions as the onset of new lactations, as many of these lactation curves resembled two contiguous lactation curves. If all observations with DIM greater than 390 days are removed from the

analysis, the results of the categorical model are not changed significantly but MthJ time in the time-based model becomes borderline significant (data not shown).

A graph of average predicted milk production compared to average observed milk production over the first 5 lactations in herd A is shown in Figure 1.1 to demonstrate that the model adequately predicts the shape of the lactation curve for the different parities; because of the similarities between cows in the older (≥2) parities in the analysis, only the first two parities were figured. Figure 1.2 shows the predicted lactation curves for different levels of infection for an average cow in herd A in the third lactation. Predicted values in both curves are generated with the time-based model (equation 1.2).

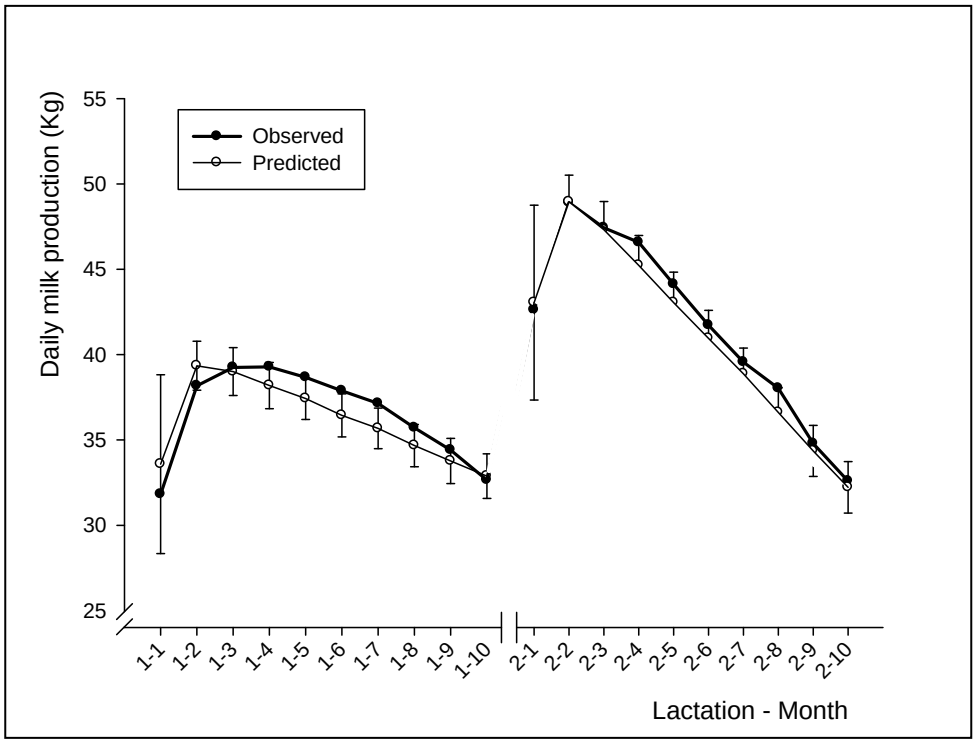



Figure 1.1: Average predicted and observed lactation curves for uninfected cows in herd A, starting in the first parity, with milk reported in kilograms (kg). The predicted values come from the full model, as detailed in Table 1.2.

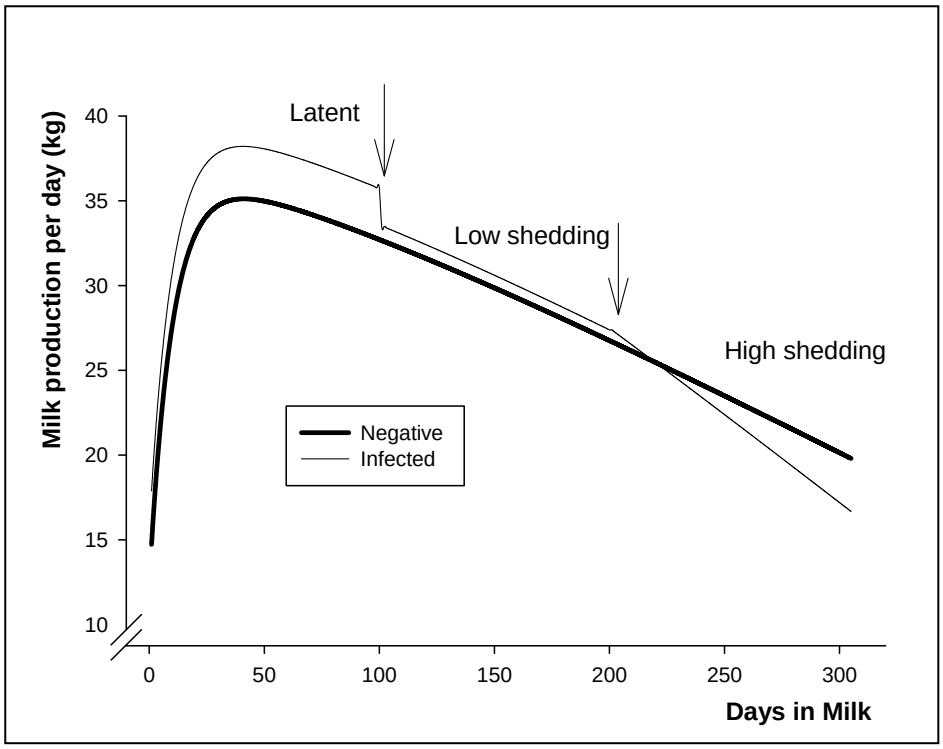



Figure 1.2: Predicted milk yield, in kilograms (kg), for an average cow in herd A in the third lactation, comparing an uninfected animal to an animal with changing infection status levels (at 100 and 200 days in milk, indicated by arrows). Changes in infection status levels were simulated for demonstrative purposes.

***Discussion***

The model presented is the first assessment of the effect of JD on milk production in individual animals over a period of several years. The results may be somewhat limited by the lack of large numbers of test day results for high-shedding animals, as few animals are retained in the herds when they are known to be shedding large quantities of MAP. Producers in the study herds were made aware of test results and were encouraged by their veterinarian, in association with the research team, to cull these animals immediately. The results may also be biased by the lack of diagnostic results for approximately 10% of animals in the study, which may have caused the loss of data on animals, especially low-producing animals, culled before

diagnosis was possible. However, the latter would have biased the results observed towards non-significance. In addition, no fecal culture results were available for 2007, which may have misclassified animals as uninfected or low-shedding. This misclassification, however, would be uninformative (assuming no correlation between milk production and presence in the herd in 2007), and should not bias the results.

All possible confounding variables included in the model (parity, DIM, LS, and calving season) were significant, with robust parameter values similar to those observed in other models of milk production (Macciotta et al., 2005), implying a large degree of stability in the analysis.

An interaction term was included to account for the unique shape of the lactation curve in first-lactation cows. Cows in their first lactation could have been modeled separately from multiparous animals. However, with the long latent period in JD, the majority of MAP-infected first-lactation cows would be latent and the latent first-lactation animals would represent a small proportion of the total MAP-infected group. This would decrease the power of our analyses, so it was therefore more valuable to combine all data in 1 model. As Figure 1.1 demonstrates, the final model was well able to predict the shape of the first lactation curve.

Johne's Disease status was found to have a significant effect on milk production, and this effect was not uniform across JD status categories. This confirms the importance of separating latent, low-shedding, and high-shedding animals in the analysis, as well as the importance of the category definitions. However, as seen in Tables 1.2 and 1.3, the effect is not always significant between groups. The difference between latent and negative or latent and low-shedding animals is relatively large and

statistically significant. The difference between negative and low-shedding animals is small and non-significant. Our observation that latent animals show an approximately 2 kg higher daily milk production compared to negative animals is currently not fully understood. However, this observation is very similar to what is observed in other infectious diseases in dairy cows. Cows with higher milk production are more susceptible to clinical mastitis (Bar et al., 2007), a relationship partially explained by a positive genetic correlation between milk production and mastitis. Similarly, in our data, cows that eventually will show low and high shedding of MAP are out-producing MAP-negative animals in the herd. Following the same argument, we speculate that there may be a genetic component to JD susceptibility, with a possible positive genetic correlation to higher milk production. This was not observed in one study examining a genetic link between ELISA values and milk production, but that study was not able to separate latent and uninfected animals in their analysis (Mortensen et al., 2004). In terms of loss of potential milk production, we would argue that MAP infected latent cows have a production potential that is 2.3 kg higher than uninfected herd mates. It is possible that low-producing latent cows were culled before infection became detectable, but the use of tissue culture at slaughter should have minimized that bias. The decrease in daily milk production due to low-shedding (2.1 kg) and high shedding (6.0 kg) of MAP would have to be accounted for in a complete economic analysis of JD. These are useful numbers for the producer to have when making economic decisions. Culling of low-shedding or high shedding animals may be based solely on the negative effect of JD on milk production, but should also take into account the contribution of shedders to MAP infection spread within the herd and other potential

negative economic effects, such as delayed reproduction. In the study herds, culling decisions were made with knowledge of MAP diagnostic results, but cattle were culled more often for low production and reproductive problems than for JD (data not shown).

In terms of this analysis, these results explain the previously observed lack of a significant difference in milk production between MAP uninfected and MAP infected animals: the higher milk production in latently MAP infected animals balances out the lower milk production in high-shedding animals (Johnson et al., 2001). Additionally, the stress associated with failing to meet the nutritional requirements of high milk production could increase the probability of shedding MAP in the future.

Considering only categorical JD status, there is a significant decrease in milk production when animals move between subsequent levels of JD status. This decrease is especially pronounced between low- and high-shedding animals. These results are not unexpected; high-shedding animals are more likely to be clinically affected by JD, with decreased performance linked to decreased intestinal absorptive capabilities (Johnson-Ifearulundu and Kaneene, 1997).

When a variable is added to the time-based model for time spent in any given JD status category, the progression of JD leads to an increase in milk production loss over time. While latent animals produce more milk than uninfected animals, that difference decreases over the course of time in the latent infection state. When an animal starts shedding low levels of MAP, the model predicts an initial milk production that is slightly higher than uninfected herd mates, but there is a greater rate of decrease in milk production compared to the latently infected animals. Finally,

animals in the high-shedding category have a meaningfully lower milk production than uninfected herd mates, with large decreases in production over time when remaining in the herd.

Figure 1.2 shows the predicted lactation curves for an average animal in herd A compared across JD status categories. A higher milk yield is evident during latency, compared to uninfected herd mates, but the discrepancy in yield decreases as the disease progresses over time. This MAP-induced decrease in milk production is supported by the clinical progression of JD. As the organism invades the intestinal epithelium and begins to affect nutrient absorption, feed efficiency decreases and milk production is negatively impacted. This is also consistent with the findings of Kudahl et al. (2004), who demonstrated that an increase in ELISA positivity (OD values) caused depressed lactation curves with more negative (decreasing) milk production slopes in late lactation. The magnitude of this effect was reported to have increased with parity (Kudahl et al., 2004). An advantage of time-based models such as the one described here (equation 1.2) is the possibility to study the continuous nature of biological processes, such as disease progression and milk production. Culling is a daily or weekly decision, and by modeling the biological process as closely as possible, the optimal time for culling any individual JD-positive animal can be narrowed to a specific month after testing results.

The main disadvantage of these models is the quality of data gathered from the available diagnostics. Sensitivity is low, especially with fecal culture, and recent studies have cast doubt on the specificity of fecal culture (Whitlock et al., 2000). Some of the animals considered JD-negative in these data may well have been latently

infected, or even intermittently shedding low quantities of MAP. Some of the low-shedding (or ELISA-positive only) animals may be experiencing "passive infections", in which they are not truly shedding MAP, but simply latently infected animals serving as passive vectors for MAP organisms present in their environment (Whitlock et al., 2008).

It was assumed that all ELISA-positive animals were low-shedding unless a high-positive fecal culture existed. This should be an acceptable assumption, as ELISA results indicative of heavy shedding have been observed to follow the corresponding fecal culture results (van Schaik et al., 2003). However, fecal culture frequency was lower than that of ELISA tests (biannual as compared to quarterly), so some animals may have been misclassified as low-shedding for several months after high-shedding began. In this case, the milk test results (with a monthly frequency) would be misclassified. It would be possible to assume a back-dated positive status, for example, a date between low-shedding and high-shedding results. However, producers will rarely perform fecal culture more than once a year, so MAP shedding results are generally available only on an annual basis.

Additionally, it was assumed that all MAP-infected animals were infected as calves. This allowed animals to be categorized as latent at all test days previous to their first positive diagnosis. The idea of calf-hood infection or, at the very latest, infection in the first year of life, is supported by a recent review article (Begg and Whittington, 2008), so this assumption was considered to be reasonable.

This analysis provides strong support that JD status impacts milk production in all infected animals, with increasing losses in milk production as disease progresses. The

above-average milk production of animals later affected by JD further highlights the important loss of milk production in high potential animals. These results will be useful in making culling decisions on an individual-animal, economic level, especially as animals shedding MAP also spread the infection through environmental contamination.

# REFERENCES

2007. NAHMS 2007: Johne's Disease on U.S. Dairy Operations. USDA:APHIS:VS:NAHMS, Fort Collins, CO, pp. 1-4.

Clarke, C.J., 1997. The pathology and pathogenesis of paratuberculosis in ruminants and other species. J. Comp. Pathol. 116, 217-261.

Dorshorst, N.C., Collins, M.T., Lombard, J.E., 2006. Decision analysis model for paratuberculosis control in commercial dairy herds. Prev. Vet. Med. 75, 92-122.

Eamens, G.J., Whittington, R.J., Marsh, I.B., Turner, M.J., Saunders, V., Kemsley, P.D., Rayward, D., 2000. Comparative sensitivity of various faecal culture methods and ELISA in dairy cattle herds with endemic Johne's disease. Vet. Microbiol. 77, 357-367.

Huda, A., Jensen, H.E., 2003. Comparison of histopathology, cultivation of tissues and rectal contents, and interferon-gamma and serum antibody responses for the diagnosis of bovine paratuberculosis. J. Comp. Pathol. 129, 259-267.

Johnson, Y.J., Kaneene, J.B., Gardiner, J.C., Lloyd, J.W., Sprecher, D.J., Coe, P.H., 2001. The effect of subclinical Mycobacterium paratuberculosis infection on milk production in Michigan dairy cows. J. Dairy Sci. 84, 2188-2194.

Kennedy, D., Benedictus, G., 2001. Control of Mycobacterium avium subsp. paratuberculosis infection in agricultural species. Rev. Sci. Technol. 20, 151-179.

Kudahl, A., Nielsen, S.S., Sorensen, J.T., 2004. Relationship between antibodies against Mycobacterium avium subsp. paratuberculosis in milk and shape of lactation curves. Prev. Vet. Med. 62, 119-134.

Lombard, J.E., Garry, F.B., McCluskey, B.J., Wagner, B.A., 2005. Risk of removal and effects on milk production associated with paratuberculosis status in dairy cows. J. Am. Vet. Med. Assoc. 227, 1975-1981.

Losinger, W.C., 2005. Economic impact of reduced milk production associated with Johne's disease on dairy operations in the USA. J. Dairy Res. 72, 425-432.

Ott, S.L., Wells, S.J., Wagner, B.A., 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. Prev. Vet. Med. 40, 179-192.

Pradhan, A., Van Kessel, J.S., Karns, J.S., Wolfgang, D.R., Hovingh, E., Nelen, K.A., Smith, J.M., Whitlock, R.H., Fyock, T.L., Ladely, S., Fedorka-Cray, P.J., Schukken, Y.H., 2009. Dynamics of endemic infectious diseases of animal and human importance on three dairy herds in the Northeastern US. J. Dairy Sci. 92, 1811-1825.

Raizman, E.A., Fetrow, J.P., Wells, S.J., Godden, S.M., Oakes, M.J., Vazquez, G., 2007. The association between Mycobacterium avium subsp. paratuberculosis fecal shedding or clinical Johne's disease and lactation performance on two Minnesota, USA dairy farms. Prev. Vet. Med. 78, 179-195.

van Schaik, G., Rossiter, C.R., Stehman, S.M., Shin, S.J., Schukken, Y.H., 2003. Longitudinal study to investigate variation in results of repeated ELISA and culture of fecal samples for Mycobacterium avium subsp paratuberculosis in commercial dairy herds. Am. J. Vet. Res. 64, 479-484.

Whitlock, R.H., Wells, S.J., Sweeney, R.W., Van Tiem, J., 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): sensitivity and specificity of each method. Vet. Microbiol. 77, 387-398.

Wilmink, J.B.M., 1987. Adjustment of test-day milk, fat, and protein yield for age, season and stage of lactation. Livest. Prod. Sci. 16, 335-348.

Wilson, D.J., Gonzalez, R.N., Hertl, J.A., Schulte, H.F., Bennett, G.J., Schukken, Y.H., Grohn, Y.T., 2004. Effect of clinical mastitis on the lactation curve: a mixed model estimation using daily milk weights. J. Dairy Sci. 87, 2073-2084.

Wilson, D.J., Rossiter, C.A., Han, H.R., Sears, P.M., 1993. Association of Mycobacterium paratuberculosis infection with reduced mastitis, but with decreased milk production and increased cull rate in clinically normal dairy cows. Am. J. Vet. Res. 54, 1851-1857.

# CHAPTER 2

# THE EFFECT OF JOHNE'S DISEASE STATUS ON REPRODUCTION AND CULLING IN DAIRY CATTLE

*originally published as:*

*The Effect of Johne's Disease Status on Reproduction and Culling in Dairy Cattle. Smith, R.L.; Strawderman, R.L.; Schukken, Y.H.; Wells, S.J.; Pradhan, A.K.; Espejo, L.A.; Whitlock, R.H.; Van Kessel, J.S.; Smith, J.M.; Wolfgang, D.R.; Grohn, Y.T. J. Dairy Sci. 2010 93(8):3513-3524.*

***Abstract***

Among the costs attributed to *Mycobacterium avium* subsp. *paratuberculosis* (MAP) infection in dairy cattle, the impacts on reproduction and culling are the least documented. In order to estimate the cost of MAP infections and Johne's Disease (JD) in a dairy herd, the rates of calving and culling were calculated for cows in each stage of MAP infection relative to uninfected cows. Data from 6 commercial dairy herds, were used for analysis, consisting of 2,818 cows with 2,754 calvings and 1,483 cullings. Every cow in each study herd was tested regularly for MAP, and herds were followed for between 4 and 7 years. An ordinal categorical variable for JD status (test-negative, low-shedding and/or ELISA-positive, or high-shedding) was defined as a time-dependent variable for all cows with at least one positive test result or two negative test results. A Cox regression model, stratified on herd and controlling for the time-dependent infection variable, was used to analyze time to culling. Non-shedding animals were significantly less likely to be culled in comparison with animals in the low-shedding/ELISA-positive category, and high-shedding animals had

non-significantly higher culling rates than low-shedding/ELISA-positive animals. Time to calving was analyzed using a proportional rates model, an analog to the Andersen-Gill regression model suitable for recurrent event data, stratifying on herd and weighted to adjust for the dependent censoring caused by the culling effects described above. High-shedding animals had lower calving rates in comparison with low-shedding/ELISA-positive animals, which tended to have higher calving rates than test-negative animals.

***Introduction***

Johne's Disease (JD) is a chronic disease of ruminants caused by intestinal infection with the pathogen *Mycobacterium avium* subsp. *paratuberculosis* (MAP). The pathogen is pervasive on US dairy farms, with approximately 68% of herds infected (USDA:APHIS:VS:NAHMS, 2007). Infection with MAP typically occurs in calves (Clarke, 1997), which, after a period of transient shedding (van Roermund et al., 2007), then enter a latent, non-shedding stage of varying length. This latent stage is followed by a period of low and intermittent shedding of MAP with no obvious clinical symptoms (Whitlock et al., 2000). If left to progress, clinical JD may develop, with a high level of MAP shedding (Whitlock et al., 2000).

Johne's Disease has been estimated to cost the US dairy industry $200 to $250 million annually (Ott et al., 1999). Two of the possible costs associated with JD are forced early culling and decreased reproductive performance. The primary economic factor in dairy production is undoubtedly milk production, but culling and reproduction are important secondary factors. Culling at the appropriate time allows for replacement of animals with higher-production-potential animals. In contrast,

culling too early can remove animals from the herd before reaching their production potential, resulting in an overall economic loss. Calving can improve milk production by initiating a new lactation, increasing milk production for a period of months. Reproduction also provides calves for sale or replacement. Both culling and reproduction may be impacted by infectious diseases (Fourichon et al., 2000; Grohn et al., 1998).

Animals with JD may be culled due to onset of clinical signs, such as diarrhea or wasting (Collins, 2003). Animals may also be culled for decreasing milk production, which is associated with progression of JD (Smith et al., 2009). In addition, structured JD control programs often recommend immediate culling of high-shedding animals to limit environmental contamination and the transmission of MAP to herd mates, particularly calves. Culling has been shown to be a potentially effective tool to reduce prevalence, but increased culling by itself is not expected to eliminate MAP infection from a dairy (Lu et al., 2008). In addition, a combination of imperfect diagnostic techniques and slow development of clinical symptoms often results in delayed culling or retention of low-shedding animals. On most farms, low-shedding animals are retained due to absence of clinical signs in these animals and the high cost of raising or purchasing replacements (Dorshorst et al., 2006).

It is hypothesized that clinical JD may cause a negative-energy and protein balance, which decreases fertility in dairy cows. However, not all studies have found an association between JD and reproductive variables, and some have found an association in the opposite direction. One prospective cohort study found that ELISA-positive animals had significantly higher numbers of days open, but that fecal culture-

positive animals had non-significantly fewer days open (Johnson-Ifearulundu et al., 2000). Several older studies found that MAP-infected cows were culled younger (Buergelt and Duncan, 1978), culled more frequently for owner-reported infertility (Merkal et al., 1975), or experienced longer calving intervals (Abbas et al., 1983) than uninfected cows. A later longitudinal study did not find an association between MAP shedding and either calving success or early culling in seasonally bred herds (de Lisle and Milestone, 1989), although analysis of infertility was not the aim of that study and the observation was made using data collected for other purposes, meaning that reproductive issues were inferred rather than directly observed. A slightly more recent case control study also found no association between JD status and calving interval (McNab et al., 1991). A recent study even found that cows with a MAP ELISA-positive serum test spent fewer days in a non-pregnant state (Lombard et al., 2005). It is important to recognize that this may be a biased result as ELISA-positive animals may be culled earlier. With these conflicting results, biases and confounding must play a major role. One study has shown that sub-clinical MAP infection, as measured by ELISA positivity, may be associated with an increase in conception rates (Marcé et al., 2009), but this association was seen to decrease with increased age and parity. Another study found that pregnancy rate increased with MAP positivity if ELISA and fecal culture were considered in parallel or if the analysis was limited to ELISA-positive animals, but that limiting the analysis to animals positive by fecal culture resulted in a non-significant decrease in pregnancy rate (Gonda et al., 2007).

Studies of reproduction need to be carefully analyzed to control for proper time at risk in all animals. In particular, when the risk of censoring (culling) is related to

both the risk of being ELISA-positive and the risk of the event (calving), care must be taken in the analyses to avoid biases due to dependent censoring. This issue of dependent censoring was highlighted in a recent prospective study. Raizman et al. (2007) evaluated the effect of MAP infection on reproduction in dairy cattle and found that infected animals are less likely to conceive (Raizman et al., 2007). Although this study employed standard time to event analysis to evaluate the impact of MAP status on reproduction, it still identified two potential biases that were not controlled. First, infected animals were bred fewer times than uninfected animals, thereby decreasing the probability of conception. Second, MAP-infected animals were culled from the herd early based on clinical JD. It has been previously shown that simulated positive dependence between culling and conception may bias the results of a proportional-hazards model for time to conception (Allore et al., 2001). Since culling plays the role of a censoring variable in statistical models for calving, bias can also occur in the analysis of calving rates unless one properly controls for the interdependence between calving rates, culling rates, and JD.

The purpose of this study was to obtain estimates of the effect of JD on both time to culling and time to calving in dairy cattle, using recent advances in statistical methods for analyzing recurrent event outcomes (i.e., calvings) subject to dependent censoring (i.e., culling). The results can be applied to economic models considering the cost of JD to the dairy industry.

***Materials and Methods***

These data are the combination of 2 datasets. The first dataset is from of a longitudinal study on three dairy herds enrolled in an ARS-Regional Dairy Quality

Management Alliance (RDQMA) study. In 2004, the RDQMA identified a commercial dairy herd in each of 3 of its member states (New York, Vermont, and Pennsylvania) to serve as longitudinal study herds. The details of this study have been described previously (Pradhan et al., 2009). Briefly, these 3 herds were visited on a quarterly basis to collect individual animal samples. During these visits all production, breeding, and health records were obtained through copying on-farm electronic records and through records obtained from the Dairy Herd Improvement Association (DHIA). Herds A and C used a synchronization program for reproduction only sporadically, and herd C bred some animals naturally. Serum samples were collected quarterly and fecal samples were collected biannually from each adult animal in each herd, and all samples were shipped overnight to The University of Pennsylvania Johne's Laboratory for analysis (Pradhan et al., 2009). Serum samples were tested by the ParaChek® (Prionics USA Inc., La Vista, NE; formerly CSL/Biocor) enzyme-linked immunosorbent assay (ELISA) for antibody reactions to MAP antigens, for which sensitivity and specificity estimates for animals shedding MAP are 0.24-0.80 and 0.98-0.99, respectively (Nielsen and Toft, 2008). Fecal samples were tested by four-tube fecal culture for presence of viable MAP organisms (Pradhan et al., 2009) and the total sum of cfu across 4 tubes was multiplied by 5.3 to determine cfu/g, with estimated sensitivity and specificity of 0.7-0.74 and 1.0, respectively, for animals shedding MAP (Nielsen and Toft, 2008). Diagnostic results were reported as positive or negative for ELISA and as cfu/g for culture. Herd owners were informed of the results of all diagnostic tests and encouraged to cull animals with high-positive fecal cultures.

The second dataset used was collected from herds in Minnesota participating in the Johne's Disease Demonstration Herd Project (JDDHP), which began in 2000 and has been described previously (Ferrouillet et al., 2009). Briefly, serum and feces were collected from all adult cows at the initiation of the study and once per year in following years, either upon confirmed pregnancy or at a single date. All monthly production, breeding, and culling records were obtained through the DHIA. For repeat breedings, Herds D and E used a synchronization program and herd F used natural breeding. Serum samples and fecal samples were processed by the Minnesota Veterinary Diagnostic Laboratory. Serum samples were tested using an ELISA test kit for serum antibody detection (IDEXX Laboratories, Inc. Westbrook, ME) as indicated by the test kit label, with estimated sensitivity and specificity for animals shedding MAP of 0.24-0.74 and 0.88-1, respectively (Nielsen and Toft, 2008). Fecal samples were tested for presence of viable MAP organisms using bacterial culture with Herrold's egg yolk (HEY) media using 72 h of sedimentation (Wells et al., 2002), with estimated sensitivity and specificity for animals shedding MAP of 0.7-0.74 and 1.0, respectively (Nielsen and Toft, 2008). Diagnostic test results were reported as positive or negative for ELISA and as cfu/tube for culture. Herd owners agreed to implement control measures that included limiting exposure of youngstock to adult cows and testing to identify and remove or quarantine infectious animals.

For all cows, Johne's Disease status was initially divided into 3 categories: test-negative, low-positive (low-shedding or ELISA-positive only), and high-shedding (defined below). At each time *t*, test-negative animals were defined as animals for which there were at least 2 diagnostic test results reported and for which all diagnostic

tests were negative. Animals with only 1 diagnostic test, if the result was negative, were removed from the analysis due to the lack of diagnostic sensitivity. Animals were assumed infected if at least 1 diagnostic test had a positive result. Animals were defined as low-positive after a positive ELISA or fecal culture of less than 30 cfu/g (RDQMA) or 50 cfu/tube (JDDHP). ELISA status has been shown to correlate with fecal shedding, allowing ELISA-positive animals to be classified as low-positive (Nielsen, 2008; van Schaik et al., 2003). Sensitivity differences between test methods were assumed to be negligible for purposes of this analysis. Animals were classified as high-shedding upon detecting a fecal culture of at least 30 cfu/g or 50 cfu/tube, representing the midpoint of a previously defined moderate shedding level (Whitlock et al., 2000). After classification as high-shedding, animals remained in this category; this is consistent with the progression of disease observed in the study herds. For purposes of analysis, we thus considered JD status to be an ordinal categorical variable, *JD(t)*, in which test-negative animals are the referent group and high-shedding animals are to be compared to low-positive animals. This ordinal structure is intended to reflect disease transitions through time (e.g., no detectable JD, progressing to subclinical JD, progressing on to clinical JD). Animals shedding high levels of MAP presumably must first pass through a low-shedding phase (whether or not such a state was observed); hence, a comparison between these two status levels is both more natural and informative than is a direct comparison between high shedding and test-negative animals.

The analyses of the relationship between time to culling and Johne's Disease can be performed using standard methods for time-to-event analysis, such as Cox's

proportional hazards model (Cox, 1972). However, because culling rates depend on other factors (e.g., parity, herd) in addition to JD status, care must be taken in order to obtain an accurate estimate of the impact of JD on calving rates. Here, we use methods described by Miloslavsky et al. (2004) to fit a proportional rates regression model that allows us to characterize the impact of JD on calving rates, a recurrent event, adjusting for dependent censoring by culling. The proportional rates regression model for recurrent events is described in Lin et al. (2000) and is directly related to the Andersen-Gill regression model (Andersen and Gill, 1982).

*Culling*

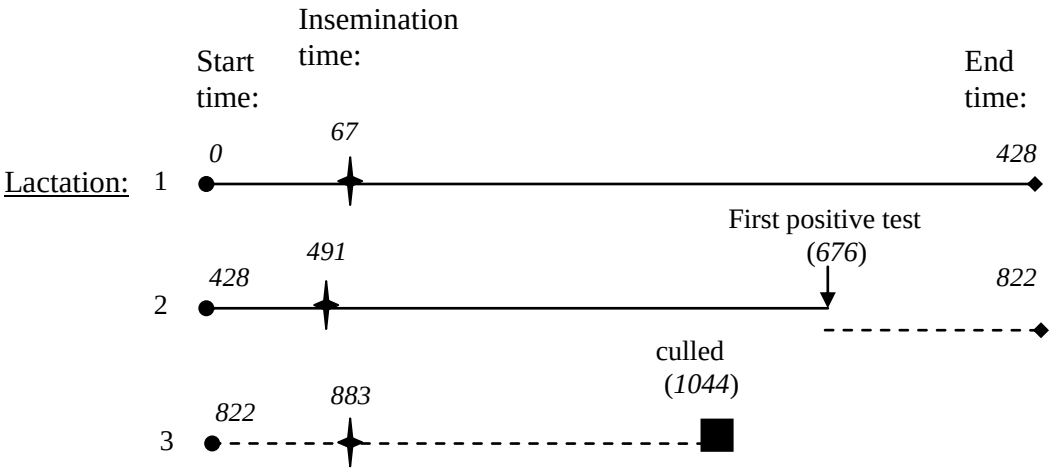
First calving and culling dates were reported for each animal by their respective herd managers. Time to culling was calculated as the number of days between first calving and culling; 22 animals were removed from the dataset because they were culled before their first calving. For each animal, this time was divided into separate records for each lactation, as shown in Figure 2.1, with time-at-risk in that parity beginning at the recorded calving date and ending with censoring at the following calving date when the animal is retained in the herd. If the animal was diagnosed antemortem, the record for the parity in which it was diagnosed was censored at the first positive test and a new record was created for the animal in the JD status category.

A Cox regression model was used to analyze time to culling; data were stratified by herd and lactation was included as a fixed effect. This regression model relates time to culling to these covariates through the following hazard function:

**Equation 2.1**

$$\lambda(t) = \lambda_{0H}(t)\exp\left(\beta_1 JD_1(t) + \beta_2 JD_2(t) + \beta_3 P_2(t) + \beta_4 P_3(t) + \beta_5 P_4(t) + \beta_6 P_5(t)\right)$$

where $\lambda(t)$ is the hazard of culling at time $t$, $t$ is the age in days, $JD_1(t)$ is a binary variable indicating that the animal is low-positive at time $t$, $JD_2(t)$ is a binary variable indicating that the animal is shedding high levels of MAP at time $t$, $P_i(t)$ is a binary




| Parity | JD status | Culling/calving model time [start end) | Culling event | Calving event | Insemination model time [start end) | Insemination model event |
|---|---|---|---|---|---|---|
| 1 | test-negative | [0 428) | 0 | 1 | [0 67) | 1 |
| 2 | test-negative | [428 676) | 0 | 0 | [428 491) | 1 |
| 2 | low-positive | [676 822) | 0 | 1 | N/A[1] | N/A[1] |
| 3 | low-positive | [822 1044) | 1 | 0 | [822 883) | 1 |

[1]Not Applicable

Figure 2.1: Diagram of datalines for the models for time to culling, calving, and insemination for a hypothetical animal. Solid circles indicate the start of a lactation, with the time in italics above being the start time for that lactation's first dataline. Stars indicate the first insemination during a given lactation. Diamonds indicate the end of a lactation, with the time in italics above. Arrows indicate positive test results, at which an animal is censored for the current dataline and a new dataline is started with the number in italics. The solid square in the third lactation indicates culling, with the time in italics above. Johne's Disease status is indicated by the line, with solid lines indicating non-shedding status and dashes indicating MAP shedding. The chart shows the data that would be used in fitting each of the model

variable indicating that an animal is in parity *i* at time *t*, $\lambda_{0H}$ is an arbitrary and unspecified baseline hazard function for the given herd, $\beta_1$ and $\beta_2$ are the parameters associated with animals in the low-positive and high-shedding JD categories, respectively, and $\beta_3$-$\beta_6$ are the parameters associated with each parity (where parity 1 is the baseline). The variable $P_5(t)$ captures all animals having a parity of 5 or more. Since nearly all censoring in this analysis occurs as a result of animals remaining under observation until the end of the study, it is reasonable to assume that censoring occurs independently of culling for the purposes of this analysis. Importantly, the effects of JD and parity are both coded as ordinal categorical variables; thus, for example, while the regression coefficient $\beta_1$ continues to measure the change in risk that occurs in moving from the test-negative referent group to the low-positive group, the regression coefficient $\beta_2$ measures the change in risk that occurs in moving from the low-positive group to the high-shedding group (i.e., as opposed to measuring the change in risk relative to the referent group). A similar interpretation applies to each of the regression coefficients for parity.

To account for a possible synergistic effect between parity and JD status, a categorical variable, *JP(t)*, was also created. This is an interaction indicating parity > 3 and positive for MAP. This resulted in the following hazard function:

**Equation 2.2**

$$\lambda(t) = \lambda_{0H}(t)\exp\begin{pmatrix}\beta_1 JD_1(t) + \beta_2 JD_2(t) + \beta_3 P_2(t) + \beta_4 P_3(t) + \beta_5 P_4(t) + \beta_6 P_5(t) \\ + \beta_7 JP(t)\end{pmatrix}$$

where *JP(t)* is a binary variable indicating that an animal is in parity 4 or higher and positive for MAP at time *t*, and $\beta_7$ is its parameter.

*Calving*

For each animal, calving dates were recorded for each lactation by the herd managers. Calculations were based on time elapsed from the first calving date (Figure 2.1). The end of the period of interest was the subsequent calving date, if available, or the date of right-censoring. If the animal was culled, the culling date was used as the end of time at-risk. If the animal was recorded as "do not breed" (**DNB**) in the herd records, the date of DNB was used as the end of time at-risk rather than the culling date; for the purposes of the reproductive model, any reference to "culling" includes DNB animals. All animals were censored at the end of the study period.

Two models were used to analyze the effect of each of several JD status categories on time to calving, a recurrent event. The first model is a proportional rates model for calving (Lin et al. 2000), with rate function given by:

**Equation 2.3**

$$\gamma_R(t) = \gamma_{0H}(t)\exp\left(\theta_1 JD_1(t) + \theta_2 JD_2(t) + \theta_3 P_2(t) + \theta_4 P_3(t) + \theta_5 P_4(t) + \theta_6 P_5(t)\right)$$

where $\gamma_R(t)$ is the risk of calving at time $t$, and $\gamma_{0H}(t)$ is an arbitrary and unspecified baseline rate function associated with herd $H$. The binary variables $JD_1(t)$ , $JD_2(t)$ , and $P_i(t)$ are each coded as in the culling model of Eq. 2, thereby continuing to represent ordinal categorical effects, and the regression coefficients $\theta_1 - \theta_6$ corresponding to the main effects therefore have interpretations analogous to those effects in the culling model, measuring the change in calving rates as an animal moves between successive (i.e., adjacent) categories. This rate function makes no explicit assumptions regarding the relationship between the current rate and the past event history beyond the dependence on parity reflected in Equation 2.3. Valid estimates of the model

parameters appearing in Equation 2.2 may be obtained in the presence of censoring due to culling provided that calving and culling rates are independent given herd, current parity and current JD status (Lin et al., 2000; Miloslavsky et al., 2004). Since JD status may be positively associated with increasing parity, the variable *JP(t)*, described above, was added to this model, giving the rate function:

**Equation 2.4**

$$\gamma_R(t) = \gamma_{0H}(t)\exp\begin{pmatrix}\theta_1 JD_1(t) + \theta_2 JD_2(t) + \theta_3 P_2(t) + \theta_4 P_3(t) \\ + \theta_5 P_4(t) + \theta_6 P_5(t) + \theta_7 JP(t)\end{pmatrix}$$

where *JP(t)* is a binary variable indicating that an animal is in parity 4 or higher and positive for MAP at time *t*, and $\theta_7$ is its parameter. It is important to note that these models each measure the impact of JD status *conditionally upon* the level of parity; in other words, the model of Equation 2.2 characterizes the impact of current JD status on calving rates after adjusting for any impact that past JD status might have on current and past parity. As illustrated in Wolfe and Strawderman (1996), care is needed to avoid fitting models that may inadvertently adjust for the effect of one time-dependent variable (i.e., JD status) through the path of another (i.e., parity).

The second model, also a proportional rates regression model, is given by

**Equation 2.5**

$$\gamma_{JD}(t) = \gamma_{1H}(t)\exp\left(\eta_1 JD_1(t) + \eta_2 JD_2(t)\right)$$

where $\gamma_{JD}(t)$ is the risk of calving at time *t*, $\eta_i$ measures the direct (i.e., unconditional) impact of the $i^{th}$ JD status $JD_i(t)$, and $\eta_{1H}(t)$ is an arbitrary and unspecified baseline rate function associated with herd *H*. The interpretation of Eq. 5 is similar to Eq. 3, describing the calving rate among animals at risk adjusting for herd and JD status

only, with JD status coded ordinally as in the models described earlier. However, unlike Eq. 3, the impact of JD status on calving rates is now considered to be unconditional, and therefore intends to measure the impact of disease regardless of the level of parity. Unfortunately, estimates obtained from this model are also problematic in the presence of censoring by culling. In particular, since culling depends on parity, the failure to control for parity when analyzing the impact of JD status on calving rates actually creates dependent censoring, leading to bias in the estimated model parameters (Miloslavsky et al., 2004).

In order to characterize the impact of JD status on reproduction rates using a model like that described in Eq. 5, adjustments for the bias induced by dependent censoring need to be made. In cases where the dependent censoring variable can be modeled, as in the case of culling, such bias can be corrected through the use of a weighting factor representing the inverse of the probability of censoring (Robins and Rotnitzky, 1992). An inverse probability of censoring weighted (**IPCW**) estimator of Miloslavsky et al. (2004) can be used for this purpose. Therefore, we also consider estimating the parameters of the model in Eq. 5 using an IPCW-based method of estimation in which animals having a high probability of censoring, such as animals that have not been successfully impregnated early in lactation, are given greater weight to correct for the relative paucity of events.

The implementation of the IPCW estimator of Miloslavsky et al. (2004) in the current context requires survivor function estimates for each animal in each lactation obtained from two different culling models. Using Equation 2.1, the first culling

model included stratification by herd and adjusts for both JD status and parity, giving a survivor function for culling of:

**Equation 2.6**

$$S_{1i}(t;R_i(t)) = \exp\left(-\int_0^t \lambda_{0H}(u)\exp\begin{pmatrix}\beta_1 JD_1(u)+\beta_2 JD_2(u)+\beta_3 P_2(u)\\ +\beta_4 P_3(u)+\beta_5 P_4(u)+\beta_6 P_5(u)\end{pmatrix}du\right)$$

where $S_{1i}(t,R_i(t))$ is the survivor function (probability of having not been culled) at time $t$, $R_i(t)$ refers to the event history for animal $i$ at time $t$, $\lambda_{0H}$ is the baseline culling rate for herd group $H$, $i$ indexes animal, and all other parameters are defined as for equation 1. Analogously to Equation 2.5, the second culling model included only stratification by herd and adjustment for JD status, giving a culling survivor function:

**Equation 2.7**

$$S_{2\,1i}(t;Z_i(t)) = \exp\left(-\int_0^t \lambda_{2H}(u)\exp\left(\delta_1 JD_1(u)+\delta_2 JD_2(u)\right)du\right)$$

where $S_{2i}(t;Z_i(t))$ is the survivor function at time $t$, $Z_i(t)$ refers to the history of the $JD$ variable for animal $i$ until time $t$, $\lambda_{2H}$ is the baseline culling rate for herd group $H$, and $\delta_1$ is the coefficient for the effect of JD status, $JD_i(t)$. The ratio $S_{2i}/S_{1i}$ was calculated for each cow-lactation-JD status combination (i.e., through time) to define a time dependent weight $w_i(t)$ that was subsequently used to estimate a culling-adjusted calving rate given MAP infection status only. The weight function $w_i(t)$ intends to adjust for the resulting impact of dependent censoring, assumed here to be created by omitting those variables from Equation 2.5 that appear in Equation 2.3. A second version of $w_i(t)$ was also created by including the interaction term $JP(t)$ in Equation 2.6, allowing us to consider the effect of the interaction between positive MAP status

and parity > 3 on culling-adjusted calving rates. Estimation of the regression parameter $\eta$ and baseline rate function $\gamma_{1H}(t)$ are easily accomplished using Cox regression software, provided this software provides the capability of incorporating these time-dependent weights. For example, the baseline rate $\gamma_{1H}(t)$ is estimated via

**Equation 2.8**

$$\hat{\gamma}_{1H}(t) = \frac{\sum_i w_i(t) dN_i(t)}{\sum_i w_i(t) Y_i(t) \exp\{\hat{\eta}_1 JD_1(t) + \hat{\eta}_2 JD_2(t)\}}$$

where $dN_i(t)$ indicates an incremental increase in the count of events, in this case a calving, $Y_i(t)$ indicates that animal $i$ is at risk for calving at time $t$ and $\hat{\eta}_i$ is the estimated coefficient for the effect of the JD status variable $JD_i(t)$. If censoring were independent of calving given only herd and JD status, $w_i(t)$ should approximately be equal to one for each $t$ and Eq. 8 collapses to the standard estimator for the baseline rate function.

As management bias may enter the analysis, and producers may exert less effort to breed animals known to be infected, time to first insemination is also a variable of interest. An unweighted Andersen Gill model with control for parity, similar to Equation 2.2, was fit to this data. Times were calculated as days elapsed from first calving, with start time being the time of the current calving and end time being the time of the first insemination for that lactation, as shown in Figure 2.1.

All survival and proportional rate regression analyses were performed with the PHREG procedure in SAS version 9.2 (© 2002-2008 by SAS Institute Inc., Cary, NC, USA.). PROC PHREG provides the ability to compute robust variance estimates

(required for use with proportional rate regression models) and time-dependent weights (required for adjusting the estimation of model parameters in a proportional rate regression models in the presence of dependent censoring). Variables were considered significant at the 0.05 level. An effect estimate of less than 0 indicates a lower risk of culling or rate of calving, respectively, while an effect estimate of greater than 0 indicates a higher risk or rate.

***Results***

Data used for the analyses are summarized in Tables 2.1 (culling) and 2.2 (time to calving). Herds in the study provided data on calvings and cullings for between 3.5 and 7 years, including data from cows born within 12 to 15 years prior to the end of the study period. Point estimates were similar when analysis was performed on each regional database (RDQMA and JDDHP) separately (data not shown), so region was not included in the analysis.

The results of the survival analyses for time to culling are presented in Table 2.3. The effect estimate for the low-positive MAP status variable, $JD_1(t)$, is significant for all analyses ($p < 0.01$), demonstrating that low-positive animals were more likely to be culled than test-negative animals, even after controlling for parity. The effect estimate for the high-shedding MAP status variable, $JD_2(t)$, was not significant, but the interaction term between JD positive status and parity > 3 was significant when added. Base culling rates were also observed to vary between herds. The decline observed in culling rates with increasing parity bears comment. In particular, due to the way in which parity is coded, a comparison is made between the indicated parity level and its preceding category. Because time is measured since the first calving date,

Table 2.1: Longitudinal data analyzed for the effect of Johne's Disease status on time to culling in dairy cattle

| Herd | State | Johne's Disease Status at culling/ end of study | Number of adult cows in study | Number of adult animals culled | Average parity at culling | Average time to culling (days) |
|---|---|---|---|---|---|---|
| A[1] | NY | test-negative | 595 | 349 | 2.8 | 1769 |
| | | low-positive[3] | 45 | 33 | 3.8 | 2285 |
| | | (ELISA positive only) | (27) | (21) | (4.4) | (2352) |
| | | high-shedding | 11 | 11 | 3.8 | 2074 |
| B[1] | PA | test-negative | 165 | 65 | 3.2 | 1942 |
| | | low-positive[3] | 6 | 4 | 4.5 | 2226 |
| | | (ELISA positive only) | (1) | (0) | (N/A[4]) | (N/A[4]) |
| | | high-shedding | 0 | 0 | N/A[4] | N/A[4] |
| C[1] | VT | test-negative | 207 | 51 | 2.8 | 1925 |
| | | low-positive[3] | 28 | 14 | 2.9 | 1969 |
| | | (ELISA positive only) | (10) | (4) | (3.3) | (2009) |
| | | high-shedding | 3 | 2 | 3.5 | 2169 |
| D[2] | MN | test-negative | 443 | 244 | 2.9 | 1872 |
| | | low-positive[3] | 97 | 66 | 3.1 | 1926 |
| | | (ELISA positive only) | (34) | (21) | (3.3) | (2001) |
| | | high-shedding | 33 | 22 | 3 | 1796 |
| E[2] | MN | test-negative | 644 | 304 | 2.7 | 1771 |
| | | low-positive[3] | 125 | 71 | 3 | 1875 |
| | | (ELISA positive only) | (36) | (24) | (3.1) | (1989) |
| | | high-shedding | 29 | 22 | 3.1 | 1776 |
| F[2] | MN | test-negative | 342 | 187 | 2.9 | 1938 |
| | | low-positive[3] | 42 | 35 | 3.2 | 2033 |
| | | (ELISA positive only) | (13) | (12) | (3.3) | (2135) |
| | | high-shedding | 3 | 3 | 1.7 | 1364 |

[1]Data from a longitudinal study on 3 dairy herds through the Regional Dairy Quality Management Alliance from 2004 to 2007, visited quarterly to collect individual animal samples and all production, breeding, and health records.

[2]Data from the Minnesota Johne's Disease Demonstration Herd Project, which provided annual individual animal samples from each herd from 2000 to 2007. All monthly production, breeding, and culling records were obtained through the DHIA.

[3] Low-positive results are defined as either fecal cultures with ≤30 cfu/g of MAP and/or positive ELISA results

[4] Not Applicable

Table 2.2: Longitudinal data analyzed for the effect of Johne's Disease status on calving interval in dairy cattle

| Herd | State | Johne's Disease Status at calving | Number of cow-lactations in study | Number of calvings | Average time of calving interval (days) |
|---|---|---|---|---|---|
| A[1] | NY | test-negative | 1758 | 990 | 402.0 |
| | | low-positive[3] | 32 | 12 | 394.3 |
| | | (ELISA positive only) | (18) | (6) | (417.0) |
| | | high-shedding | 7 | 1 | 391.0 |
| B[1] | PA | test-negative | 239 | 74 | 407.8 |
| | | low-positive[3] | 2 | 0 | N/A[4] |
| | | (ELISA positive only) | (0) | (0) | (N/A[4]) |
| | | high-shedding | 0 | N/A[4] | N/A[4] |
| C[1] | VT | test-negative | 406 | 171 | 427.1 |
| | | low-positive[3] | 21 | 5 | 490.0 |
| | | (ELISA positive only) | (3) | (0) | (N/A[4]) |
| | | high-shedding | 4 | 1 | 362.0 |
| D[2] | MN | test-negative | 986 | 454 | 446.9 |
| | | low-positive[3] | 196 | 92 | 406.7 |
| | | (ELISA positive only) | (31) | (18) | (399.5) |
| | | high-shedding | 10 | 3 | 380.3 |
| E[2] | MN | test-negative | 1352 | 624 | 412.7 |
| | | low-positive[3] | 82 | 10 | 395.6 |
| | | (ELISA positive only) | (21) | (4) | (378.5) |
| | | high-shedding | 12 | 0 | N/A[4] |
| F[2] | MN | test-negative | 705 | 310 | 434.5 |
| | | low-positive[3] | 32 | 7 | 412.7 |
| | | (ELISA positive only) | (7) | (2) | (414.0) |
| | | high-shedding | 1 | 0 | N/A[4] |

[1]Data from a longitudinal study on 3 dairy herds through the Regional Dairy Quality Management Alliance from 2004 to 2007, visited quarterly to collect individual animal samples and all production, breeding, and health records.

[2]Data from the Minnesota Johne's Disease Demonstration Herd Project, which provided annual individual animal samples from each herd from 2000 to 2007. All monthly production, breeding, and culling records were obtained through the DHIA.

[3] Low-positive results are defined as either fecal cultures with ≤30 cfu/g of MAP and/or positive ELISA results

[4] Not Applicable

Table 2.3: Results of a proportional hazards model for effect of Johne's Disease (JD) status on time to culling. The model takes into account censoring due to end of the study, controls for parity, and is stratified by herd. The model was fit with and without a term for the interaction between JD status and parity > 3. 6163 observations were used in the analysis.

| Variable | Level | Effect Estimate (std. error) | |
|---|---|---|---|
| | | No Interaction | Interaction |
| Parity (1 is the base) | 2 | -1.04 (0.06) | -1.03 (0.06) |
| | 3 | -0.36 (0.06) | -0.36 (0.06) |
| | 4 | -0.09 (0.08)* | -0.14 (0.08)* |
| | 5+ | -0.27 (0.13) | -0.29 (0.11) |
| Johne's Disease Status | low-positive (base is test-negative) | 0.34 (0.06) | 0.27 (0.13) |
| | high-shedding (base is low-positive) | 0.12 (0.13)* | 0.15 (0.13)* |
| Parity*Johne's Disease Status | positive, parity>3 | N/A[1] | 0.26 (0.12) |

[1]Not Applicable
*Not significant at the 5% level.

rather than since the beginning of the lactation, animals with higher parity tend to be under observation longer, hence have longer culling times. Culling reasons were provided by herds A, B, and C for 393, 69, and 67 animals, respectively; 143 animals were recorded as having been culled for reproductive reasons, and a further 143 animals were recorded as having been culled for disease reasons, including 8 high-shedding animals recorded as having been culled for clinical JD. Other reasons were given for culling a total of 243 animals, including injuries and production.

The results for the proportional rate regression models for time to calving described in the previous section are presented in Table 2.4. For the models corresponding to Equations 2.3 and 2.4, the rate of calving decreased as parity increased (results not shown), with no evidence of a differential effect of JD status for

Table 2.4: Effect estimates are shown from 4 separate sets of proportional rates models for the effect of Johne's Disease (JD) status on calving interval, with standard error in parentheses. The unweighted model for Equation 2.3 adjusts for JD status, parity, and herd (through stratification). The unweighted model for Equation 2.4 adjusts for JD status, parity, and herd (through stratification), and the interaction between test-positive status and parity >3. The unweighted model for Equation 2.5, using the baseline function of Equation 2.8 with weights set equal one, adjusts only for JD status and herd (through stratification), dropping parity. The weighted model for Equation 2.5, using the baseline function of Equation 2.8, controls for both the probability of culling and parity through weighting, with or without the interaction between test-positive status and parity > 3, controlling for herd through stratification. The results of an unweighted proportional rates model based on time to first insemination, controlling for parity and stratified by herd, are also included. 6163 observations were included in the analysis.

| Comparison | Unweighted Effect Estimate for Time to Calving (Equation 2. 3) | Unweighted Effect Estimate for Time to Calving with Interaction (Equation 2.4) | Unweighted Effect Estimate for Time to Calving (Equations 2.5, 2.8) |
|---|---|---|---|
| Low-positive (test-negative is base) | 0.18 (0.07) | 0.17 (0.08) | 0.14 (0.06) |
| High-shedding (low-positive is base | -0.46* (0.29) | -0.45* (0.29) | -0.36* (0.23) |

| Comparison | Weighted Effect Estimate for Time to Calving (Equations 2.5, 2.8) | Weighted Effect Estimate for Time to Calving with Interaction (Equations 2.5, 2.8) | Effect Estimate for time to first insemination |
|---|---|---|---|
| Low-positive (test-negative is base) | 0.20* (0.12) | 0.23 (0.12) | 0.01* (0.06) |
| High-shedding (low-positive is base | -1.11 (0.52) | -1.07 (0.50) | -0.02* (0.13) |

*Not significant at the 5% level

lower versus higher parities. Animals low-positive for MAP (first row of Table 2.4) were also estimated to have a higher rate of calving compared to test-negative animals. This latter result was also reflected in the results obtained for both the unweighted and

weighted model fits corresponding to Equation 2.5, which no longer directly adjusted for parity in the rate model itself. The effect of low-positive status is not statistically significant ($p = 0.116$) when an interaction between parity and JD status is not included in the culling model used to compute the weights, but is borderline significant in the presence of this interaction ($p = 0.046$).

Animals shedding high levels of MAP (the second row of Table 2.4) exhibited substantial, and more disparate, numerical decreases in the calving rate compared to low-positive animals in each case, with the impact under the weighted model fit corresponding to Equation 2.5 being both the strongest and the most statistically significant. Comparison of the effect estimates for the weighted and unweighted fits corresponding to Equation 2.5 show that the latter is more than twice the former. In our view, these results clearly illustrate two features of the proposed methodology. First, a comparison of the differences in effect estimates obtained for the unweighted and weighted fits corresponding to the model of Equation 2.5 demonstrates the impact of dependent censoring that results from a failure to control for the impact that parity has on culling (i.e., the dependent censoring variable). Here, we further observe that inclusion of the interaction term *JP(t)* did not qualitatively or quantitatively affect the estimates for any model. Second, a comparison of the model fits for Equation 2.4 and the weighted fits corresponding to Equation 2.5 illustrates the differences in interpretation of the regression coefficients in models that directly, versus indirectly, adjust for parity when analyzing the impact of JD status (Wolfe and Strawderman, 1996). In particular, the regression coefficient in Equation 2.4 is attenuated, a reflection of the fact that a model that directly adjusts for both parity and JD status is

capturing the effect of JD status on calving rates that remains after adjusting for any effect it may have had on past and current parity.

Results of the analysis for time to first insemination (Table 2.4) further showed that time to first insemination was not significantly related to JD status.

***Discussion***

The models presented above analyzed the relationship between JD status (as defined by MAP shedding and ELISA results) and both the risk of culling and calving rates. The results of the model for time to culling were easily influenced by producer decisions. In all herds, producers were interested in controlling or eradicating JD and were informed of test results when available. Based on culling reasons given for high-shedding animals, in herds A and C, JD status was the most common reason stated. Thus, the results likely show a combination of both biological and producer-decision effects, with the latter influenced by expert advice.

Beyond the desire to control the spread of JD, producers often cull animals due to low or decreased milk production. The authors have previously observed, with the same dataset, a significant relationship between JD status and decreasing milk production (Smith et al., 2009). While low-positive and test-negative animals did not have significantly different average daily milk production levels in that analysis, milk production in low-positive animals did decrease over time faster than test-negative controls. High-shedding animals, when compared to all other animals, were observed to have both a significantly lower average daily milk production and a significantly greater monthly decrease in milk production. This decreased milk production observed in MAP-positive animals would be expected to lead to the increased rate of

culling in MAP-positive animals observed in the present study, as an intervening variable on the causal pathway between JD and culling. The increased culling rate could also be related to other effects of JD (including the hypothesized decreased calving rates), or to the producer's desire to control JD through culling. The former has been observed previously; fecal culture-positive cows were observed to be culled for infertility at a higher rate than culture-negative cows (Merkal et al., 1975). In these data, the latter was also known to play a role, as the herds participate in a JD control program that recommends culling of high-positive animals.

Inclusion of a term to capture the interaction between JD status and age (via the proxy of parity) was found to be necessary. However, the use of a saturated interaction model would have required 8 additional terms in the full model. In addition to creating serious challenges in model interpretation (particularly given our use of an ordinal coding strategy), the presence of relatively sparse data available for high-shedding animals raised questions of feasibility. Our decision to utilize a simpler interaction term that differentiated older, positive animals from the rest was motivated by a preliminary data analysis using a simple 4-level variable that classified animals according to JD status ("any" vs. not) and age ("parity > 3" vs. not) and which suggested that the strongest effect was in these animals (results not shown). While this interaction term was found to be statistically significant in the full culling model of Eqn. 3, its inclusion did not qualitatively change the results of the culling model or the weighted calving models.

Due to the positive dependency between calving, culling, parity and JD status, indicated by the discrepancy in the results presented in Table 4 and the large

proportion of animals culled for reasons of reproductive problems, a correction to the standard proportional rates model was required to more appropriately analyze the effect of JD status on reproduction. With the IPCWE approach of Miloslavsky et al. (2004), a direct estimate of the marginal effect of JD status can be obtained. Other methods, such as frailty models (Rondeau et al., 2007), might have been used instead to deal with the possibility of dependent censoring, as well as the dependence between multiple calvings in the same animal; however, such methods may yield less interpretable estimates of the effect of JD status. In this study, the difference between the IPCWE-corrected (i.e., weighted) and basic (i.e., unweighted) models was found to be relatively small for the case of all low-positive animals and substantially larger for high shedding animals. For both levels of JD status, the weighted analysis was also found to be more conservative, resulting in larger standard errors. The addition of an interaction term to capture the synergy between age and test-positive status did not qualitatively change the results of any model. Thus, for example, one can conclude that a synergistic relationship between age and JD status on culling rates is unlikely to be responsible for the observed increase in calving intervals in high shedding cows obtained when comparing the weighted and unweighted model fits corresponding to Equation 2.5.

There were possible biases in this study introduced by the method of measurement available. As reproduction was measured based on the proxy of calving intervals, rather than time to conception, some animals censored for the end of study may have conceived, but not reached calving. Animals were also considered at-risk for calving, in this model, during the voluntary waiting period, breeding, and early

gestation. That was not precisely true, as animals were only at risk of calving in the end of gestation. It was assumed that no bias arose from ignoring end-of-study censoring, the voluntary waiting period, and gestation time; all these should have been unassociated with JD status, and should not have impacted the estimates. This left only the effects of JD status on breeding, which included both producer decisions and biological effects on conception. However, a change in JD status may occur between conception and calving; in this analysis, the event of interest would be falsely attributed to the higher JD status. Future studies may avoid these issues by recording conception dates when collecting data, allowing for a more direct analysis. It would also be of interest to consider the effect of latent infection on production parameters, but such an analysis is not possible within the scope of these data.

The categorization of JD status could also lead to biased estimates; animals were assumed to be truly positive or negative, despite the imperfect specificity and sensitivity of the diagnostic tests. Any passive shedders (animals with positive fecal samples that were not infected) would confuse the categories used, making differences difficult to detect. False-positive results would also bias the estimates towards the null. Further analysis of the question within research herds, with tissue culture of animals on slaughter to ensure proper categorization, may lead to stronger or more significant findings.

This study also combined two separate databases, which were collected for similar purposes but with different methods. The point estimates for the effects of JD status on both culling and reproduction were similar if the two databases were

analyzed separately; thus, we concluded that combining the datasets, leading to improved precision of all estimates, was appropriate.

Milk production level is known to influence calving rates (Lucy, 2001), and the prevailing understanding is that high milk production is correlated with decreased calving rates. We did not include this in our analysis, however, in order to capture the full effect of JD status on calving rates. In our study, high shedding cows tended to have lower calving rates than others, but previous analysis of the RDQMA database showed that high shedding cows had significantly lower milk production compared to all other cows (Smith, 2009). Therefore, milk production could be considered an intervening variable in the pathway between JD status and calving rates, rather than a confounding variable. Calving season may also impact on reproduction, but calving season was found not to be associated with JD status in our data, so we did not include it in the model.

Our data also showed that low-positive animals had higher calving rates than test-negative animals. This agrees with an earlier published study (Lombard et al., 2005), which found that animals with a strong-positive ELISA had significantly fewer days open than ELISA-negative animals. However, Lombard et al. were not able to distinguish between stages of JD, as fecal culture results were not provided in that study. Marcé et al. (2009) also found higher calving rates among ELISA-positive animals, but the decrease of the effect with increasing parity suggests that disease progression may reverse the effect. Our results suggest that movement from low-positive to high shedding of MAP was associated with decreased calving rates, which would agree with Marcé et al.'s findings. The opposing effects of JD in low-positive

and high-shedding animals, also observed to some extent by Gonda et al. (2007) may explain the failure of previous studies (de Lisle and Milestone, 1989; McNab et al., 1991) to find an effect in positive animals, especially as controlling for the effect of JD on culling proved to be necessary.

Decreased calving rates can be an expensive loss for dairy herds, but the cause is uncertain. They could possibly indicate a biological cause, such as the postulated effect of clinical JD leading to a negative energy balance, which can lead to an anovulatory state (Butler et al., 2006), significantly delaying time to conception (Butler, 2003). However, it was not possible, with these data, to determine that the effect of the disease on calving was due to biological effects of disease rather than producer decisions. Not every herd in this study recorded when an animal was labeled DNB, and even herds that did record it may have done so sporadically. As such, animals that were not at risk for conception would have been mistakenly included in the dataset. Producers may also have bred MAP-positive animals less frequently than MAP-negative animals, with an unconscious or unrecorded bias, artificially decreasing their risk of conception. Raizman et al. (2007) found no association between JD status and days open, but did find that non-pregnant animals with positive fecal cultures were bred significantly fewer times than test negative animals. In the same way, animals that were high milk producers (such as latently infected animals, in this model) may have been bred more intensively, with more inseminations than lower-producing animals. For this reason, we examined the time to first insemination as a proxy of producer effort. However, time to first insemination was not significantly

affected by JD status, which suggests a biological cause for the observed effect on calving intervals.

In conclusion, survival analysis indicated that detectable infection with MAP resulted in increased culling rates. We also observed an increased calving interval in animals shedding high levels of MAP compared to low-positive animals.

# REFERENCES

2007. NAHMS 2007: Johne's Disease on U.S. Dairy Operations. USDA:APHIS:VS:NAHMS, Fort Collins, CO, pp. 1-4.

Abbas, B., Rienmann, H.P., Hird, D.W., 1983. Diagnosis of Johne's disease (paratuberculosis) in northern California cattle and a note on its economic significance. Calif. Vet. 8, 19-24.

Allore, H.G., Warnick, L.D., Hertl, J., Grohn, Y.T., 2001. Censoring in survival analysis: a simulation study of the effect of milk yield on conception. Prev. Vet. Med. 49, 223-234.

Andersen, P.K., Gill, R.D., 1982. Cox's Regression Model Counting Process: A Large Sample Study. Ann. Statist. 10, 1100-1120.

Buergelt, C.D., Duncan, J.R., 1978. Age and milk production data of cattle culled from a dairy herd with paratuberculosis. J. Am. Vet. Med. Assoc. 173, 478-480.

Butler, S.T., Pelton, S.H., Butler, W.R., 2006. Energy balance, metabolic status, and the first postpartum ovarian follicle wave in cows administered propylene glycol. J. Dairy Sci. 89, 2938-2951.

Butler, W.R., 2003. Energy balance relationships with follicular development, ovulation and fertility in postpartum dairy cows. Livest. Prod. Sci. 83, 211-218.

Clarke, C.J., 1997. The pathology and pathogenesis of paratuberculosis in ruminants and other species. J. Comp. Pathol. 116, 217-261.

Collins, M.T., 2003. Paratuberculosis: review of present knowledge. Acta Vet. Scand. 44, 217-221.

Cox, D.R., 1972. Regression models and life-tables (with discussion). J. R. Stat. Soc. Ser. B 34, 187-220.

de Lisle, G.W., Milestone, B.A., 1989. The economic impact of Johne's disease in New Zealand. In: Milner, A.R., Wood, P.R. (Eds.), Johne's Disease, Current Trends in Research Diagnosis and Management. CSIRO Publications, Melbourne, pp. 41-45.

Dorshorst, N.C., Collins, M.T., Lombard, J.E., 2006. Decision analysis model for paratuberculosis control in commercial dairy herds. Prev. Vet. Med. 75, 92-122.

Ferrouillet, C., Wells, S.J., Hartmann, W.L., Godden, S.M., Carrier, J., 2009. Decrease of Johne's disease prevalence and incidence in six Minnesota, USA, dairy cattle herds on a long-term management program. Prev. Vet. Med. 88, 128-137.

Fourichon, C., Seegers, H., Malher, X., 2000. Effects of disease on reproduction in the dairy cow: a meta-analysis. Therio. 53, 1729-1759.

Gonda, M.G., Chang, Y.M., Shook, G.E., Collins, M.T., Kirkpatrick, B.W., 2007. Effect of Mycobacterium paratuberculosis infection on production, reproduction, and health traits in US Holsteins. Prev. Vet. Med. 80, 103-119.

Grohn, Y.T., Eicker, S.W., Ducrocq, V., Hertl, J.A., 1998. Effect of diseases on the culling of Holstein dairy cows in New York State. J. Dairy Sci. 81, 966-978.
Johnson-Ifearulundu, Y.J., Kaneene, J.B., Sprecher, D.J., Gardiner, J.C., Lloyd, J.W., 2000. The effect of subclinical Mycobacterium paratuberculosis infection on days open in Michigan, USA, dairy cows. Prev. Vet. Med. 46, 171-181.

Lin, D.Y., Wei, L.J., Yang, I., Ying, Z., 2000. Semiparametric regression for the mean and rate functions of recurrent events. J. R. Stat. Soc. Ser. B 52, 711-730.

Lombard, J.E., Garry, F.B., McCluskey, B.J., Wagner, B.A., 2005. Risk of removal and effects on milk production associated with paratuberculosis status in dairy cows. J. Am. Vet. Med. Assoc. 227, 1975-1981.

Lu, Z., Mitchell, R.M., Smith, R.L., Van Kessel, J.S., Chapagain, P.P., Schukken, Y.H., Grohn, Y.T., 2008. The importance of culling in Johne's Disease control. J. Theo. Bio. 254, 135-146.

Lucy, M.C., 2001. Reproductive loss in high-producting dairy cattle: Where will it end? J. Dairy Sci. 84, 1277-1293.

Marcé, C., Beaudeau, F., Bareille, N., Seegers, H., Fourichon, C., 2009. Higher non-return rate associated with Mycobacterium avium subspecies paratuberculosis infection at early stage in Holstein dairy cows. Theriogenology 71, 807-816.

McNab, W.B., Meek, A.H., Martin, S.W., Duncan, J.R., 1991. Associations between dairy production indices and lipoarabinomannan enzyme-immunoassay results for paratuberculosis. Can. J. Vet. Res. 55, 356-361.

Merkal, R.S., Larsen, A.B., Booth, G.D., 1975. Analysis of the effects of inapparent bovine paratuberculosis. Am. J. Vet. Res. 36, 837-838.

Miloslavsky, M., Keles, S., van der Laan, M.J., 2004. Recurrent events analysis in the presence of time-dependent covariates and dependent censoring. J. R. Stat. Soc. Ser. B 66, 239-257.

Nielsen, S.S., 2008. Transitions in diagnostic tests used for detection of Mycobacterium avium subsp. paratuberculosis infections in cattle. Vet. Microbiol. 132, 274-282.

Nielsen, S.S., Toft, N., 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-g assay, and faecal culture techniques. Vet. Microbiol. 129, 217-235.

Ott, S.L., Wells, S.J., Wagner, B.A., 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. Prev. Vet. Med. 40, 179-192.

Pradhan, A., Van Kessel, J.S., Karns, J.S., Wolfgang, D.R., Hovingh, E., Nelen, K.A., Smith, J.M., Whitlock, R.H., Fyock, T.L., Ladely, S., Fedorka-Cray, P.J., Schukken, Y.H., 2009. Dynamics of endemic infectious diseases of animal and human importance on three dairy herds in the Northeastern US. J. Dairy Sci. 92, 1811-1825.

Raizman, E.A., Fetrow, J.P., Wells, S.J., Godden, S.M., Oakes, M.J., Vazquez, G., 2007. The association between Mycobacterium avium subsp. paratuberculosis fecal shedding or clinical Johne's disease and lactation performance on two Minnesota, USA dairy farms. Prev. Vet. Med. 78, 179-195.

Robins, J.R., Rotnitzky, A., 1992. Recovery of information and adjustment for dependent censoring using surrogate markers. AIDS Epidemiology: Methodological Issues. Birkhauser, Boston, pp. 24-33.

Rondeau, V., Mathoulin-Pelissier, S., Jacqmin-Gadda, H., Brouste, V., Soubeyran, P., 2007. Joint frailty models for recurring events and death using maximum penalized likelihood estimation: application on cancer events. Biostatistics 8, 708-721.

Smith, R.L., Grohn, Y.T., Pradhan, A.K., Whitlock, R.H., Van Kessel, J.S., Smith, J.M., Wolfgang, D.R., Schukken, Y.H., 2009. A longitudinal study for the impact of Johne's Disease status on milk production in individual cows. J. Dairy Sci. 92, 2653-2661.

van Roermund, H.J.W., Bakker, D., Willemsen, P.T.J., de Jong, M.C.M., 2007. Horizontal transmission of Mycobacterium avium subsp. paratuberculosis in cattle in an experimental setting: Calves can transmit the infection to other calves. Vet. Microbiol. 122, 270-279.

van Schaik, G., Rossiter, C.R., Stehman, S.M., Shin, S.J., Schukken, Y.H., 2003. Longitudinal study to investigate variation in results of repeated ELISA and culture of fecal samples for Mycobacterium avium subsp paratuberculosis in commercial dairy herds. Am. J. Vet. Res. 64, 479-484.

Wells, S.J., Wagner, B.A., Lindeman, C.J., Fyock, T.L., 2002. Evaluation of bacteriologic culture of pooled faecal samples for detection of Mycobacterium paratuberculosis. Am. J. Vet. Res. 63, 1207-1211.

Whitlock, R.H., Wells, S.J., Sweeney, R.W., Van, T.J., 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): sensitivity and specificity of each method. Vet. Microbiol. 77, 387-398.

Wolfe, R.A., Strawderman, R.L., 1996. Logical and statistical fallacies in the use of Cox regression models. Am. J. Kidney Dis. 27, 124-129.

# CHAPTER 3

# ENVIRONMENTAL CONTAMINATION WITH *MYCOBACTERIUM AVIUM* SUBSP. *PARATUBERCULOSIS* IN ENDEMICALLY INFECTED DAIRY HERDS

*In press as:*

*Environmental contamination with Mycobacterium avium subsp. paratuberculosis in endemically infected dairy herds. R. L. Smith, Y. H. Schukken, A. K. Pradhan, J. M. Smith, R. H. Whitlock, J. S. Van Kessel, D. R. Wolfgang and Y. T. Grohn. Preventive Veterinary Medicine 2011.*

***Abstract***

Environmental contamination with *Mycobacterium avium* subsp. *paratuberculosis* (MAP) is thought to be one of the primary sources of infection for dairy cattle. The exact link between fecal shedding of MAP by individual cows and environmental contamination levels at the herd level was explored with a cross-sectional analysis of longitudinally collected samples on 3 dairy farms. Composite samples from multiple environmental sites in 3 commercial dairy herds in the Northeast US were cultured quarterly for MAP, providing 1131 samples (133 (11.8%) were culture-positive), and all adult animals in the herds were tested biannually by fecal culture (FC), for 6 years. Of the environmental sites sampled, manure storage areas and shared alleyways were most likely to be culture-positive. Environmental sample results were compared to FC results from either the concurrent or previous sampling date at both the herd and the pen level. At the herd level, a 1 log unit increase in average fecal shedding increased the odds of a positive non-pen environmental sample by a factor of 6 and increased the average amount of MAP in non-pen samples by 2.9 cfu/g. At the pen level, a 1 log

unit increase in average fecal shedding in the pen increased the odds of a positive environment by a factor of 2.4 and the average amount of MAP was increased by 3.5 cfu/g. We were not able to model the relationship between non-pen environmental sample status and the distance between shedding animals and the sample's location, and neighboring pens did not significantly affect the results of the pen-level analysis. The amount of MAP in pen-level samples and the probability of a pen testing positive for MAP were both positively but non-significantly correlated with the number of animals in the pen shedding >30 cfu/g of MAP. At least 6 environmental samples met the criteria for the U.S. Voluntary Bovine Johne's Disease Control Program on 47 of the 72 sampling dates; of these, 19 of the 47 FC-positive sampling dates were positive by the 6-sample environmental testing method, resulting in a herd sensitivity of 0.40 (95% CI: 0.26 to 0.54). None of the 3 FC-negative sampling dates produced positive environmental samples. Although environmental sampling can be used as a tool in understanding the level of MAP infection in a herd or pen, it did not appear to be a sensitive diagnostic method for herd positivity in these low prevalence herds, and its use may require caution.

***Introduction***

Johne's disease in cattle is caused by a chronic intestinal infection with Mycobacterium avium subsp. paratuberculosis (MAP), due to ingestion of the organism. After a long latent period, animals infected with MAP begin to shed the organism in their feces (Benedictus et al., 2008), thereby contaminating the farm environment. This environmental contamination with MAP is thought to be one of the primary sources of infection for dairy cattle (Nielsen and Toft, 2009).

Environmental sampling has been evaluated for identification of herd MAP status (Lombard et al., 2006; Raizman et al., 2004). The number of MAP-positive environmental cultures (EC) on a dairy farm has been found to be proportionate to the seroprevalence in the herd (Berghaus et al., 2006) as well as the fecal culture (FC) prevalence (Pillars et al., 2009). In assessing the reliability of repeat environmental samples, the majority of variation in MAP concentration has appeared to come from the source of the sample, both the dairy herd and the pen within the herd (Aly et al., 2009; Pillars et al., 2009). The National Animal Health Monitoring System (NAHMS) has used standardized environmental sampling (USDA:APHIS:VS, 2010) to determine the apparent herd-level prevalence of MAP nationally, currently estimated in dairy herds at 68% (USDA:APHIS:VS, 2008). However, no studies have examined the relationship between the concentration of MAP in environmental samples and in the feces of individual cows.

The objective of this study was to longitudinally describe environmental MAP contamination in endemically infected dairy herds, and to correlate that contamination to fecal shedding by individual animals.

***Materials and Methods***

Sample collection and isolation of MAP from samples used in this study have been previously described (Pradhan et al., 2009). Briefly, 1 dairy herd in each of 3 states (herd A in New York, herd B in Pennsylvania, and herd C in Vermont) was visited quarterly by members of the Regional Dairy Quality Management Alliance (RDQMA) from 2004 through 2009. A number of environmental sites were consistently sampled on each quarterly visit from mapped locations using the same

method. Source water, water from drinking troughs of cows and heifers, and standing water or manure slurry in freestall barns were collected by a tube scooped through the water/slurry. Feed from pens of cows and calves, calf bedding, corn silage, and manure composites from milking cows, nonlactating cows, periparturient cows, calf pens, and milking alleyways were collected manually by a freshly gloved hand; at each of these sites (i.e., an alleyway), material from 4-6 locations (i.e., 4-6 points along the alleyway) was combined into one composite sample. On some sampling dates, bird droppings, flies and insects were also collected. The samples were classified as Alleyway, Pen (consisting only of adult cow pens), Pit (consisting of manure storage areas), and Other. Biannually, fecal samples were collected rectally from every adult cow present in the herds. All fecal samples were tested by four-tube fecal culture (Pradhan et al., 2009), and the total sum of cfu across 4 tubes was multiplied by the conversion factor, 5.3, to determine cfu/g (Pradhan et al., 2011).

Environments sampled included the 6 minimum sites for herd-level MAP testing recommended by the Voluntary Bovine Johne's Disease Control Program (VBJDCP) (USDA:APHIS:VS, 2010): 2 each of cow housing alleyways or gutters, manure storage areas, and another manure concentration area. The culture results of these 6 qualifying samples were combined to create a single variable, the Standard 6 (S6). When more than 6 samples qualifying for the S6 were collected, 6 were chosen randomly from the qualifying samples to include in the S6 results.

*Statistical Analysis*

Environmental samples were collected quarterly and fecal samples were collected biannually, so EC results were modeled with concurrent FC data if possible;

if no concurrent FC data were available, EC results were modeled using FC data from the previous quarter. Physical distances between sampling locations were measured as straight lines between the center of each pen and the center of the sampling location using the ruler tool in Google Earth (©2010 Google), to represent the average distance that MAP must move from the pen to the sampling location.

The vector of EC results can be defined as **Y**, where the $Y_i$ are independent Bernoulli random variables where 1 indicates the presence of MAP in an EC, $E[Y_i]=\pi(\mathbf{X_i})$ and $Var(Y_i)=\pi(\mathbf{X_i})*(1-\pi(\mathbf{X_i}))$, and $\mathbf{X_i}$ is the vector of explanatory variables for sample *i*, then

**Equation 3.1**

$$\pi(X_i) = {exp(\boldsymbol{\beta}'\boldsymbol{X_i} + \gamma' z_i)} \Big/ {1 + exp(\boldsymbol{\beta}'\boldsymbol{X_i} + \gamma' z_i)}$$

where $\boldsymbol{\beta}$ is the vector of the effects of explanatory variables, $z_i$ is the sampling date, and $\gamma$ is normally distributed with a mean of 0 and variance of $\sigma_\gamma^2$. A random effects model was used as ECs from the same sampling date were assumed to not be independent. For this logistic model, several configurations of $\mathbf{X_i}$ were considered: all contained sample type (alley, pen, pit, or other) and herd (A, B, or C), the prevalence-based model contained the natural log of the proportion of MAP-positive animals in the herd, the amount-based model contained the base-10 log of the average shedding level (cfu/g) of all individual adults, and the high-shedder-based model contained the number of animals shedding ≥30 cfu/g in the herd. In addition, the results of EC may be modeled as a continuous variable $\varphi_i$, the amount of MAP (cfu/g) in the sample. In this case,

**Equation 3.2**

$$log(\varphi_i) = \boldsymbol{\beta}'\boldsymbol{X_i} + \gamma' z_i + e_i$$

with $e_i$ being the random error associated with sample *i* and all other variables defined as above for Equation 1.

The logistic model in Equation 3.1 was also fit to the database with pen samples excluded, and to data limited to pen samples in which the FC results were limited to animals in the pen from which the sample was taken. In the model of pen samples only, sample type was removed from $\mathbf{X_i}$. The linear model in Equation 3.2 was also used for the separate databases (pen and non-pen samples), adding an additional term to $\mathbf{X_i}$ for the distance-corrected average shedding level in the pens $\sum_i {log(f_i)}/{d_i}$, where $f_i$ is the shedding level in pen *i* and $d_i$ is the distance between pen *i* and the center of the sampled area (in meters); a similar analysis was also performed using the squared distance between pens, $d^2$ ($m^2$), as the effect of distance is not necessarily linear.

*VBJDCP samples*

The primary analysis of the S6 evaluated whether there was a relationship between the presence of MAP in FC and the presence of MAP in the S6. This evaluation was based on a logistic regression model (Equation 3.1), where $\pi(\boldsymbol{X}_i)$ was the probability that any FC were MAP positive, $z_i$ was unity, and $\boldsymbol{X_i}$ consisted of the herd (A, B, or C) and either the overall result of VBJDCP sampling (positive if ≥1/6 S6 cultures was positive), the number of S6 cultures testing positive, or the base-10 log of the average amount of MAP in the 6 S6 samples (cfu/g). In any of these cases,

exp($\beta_n$) was the odds ratio for herd shedding status with a 1-unit increase in the S6 results, where $\beta_n$ was the parameter estimate associated with that result. The relationship between the results of the S6 and the MAP infection level within the herd was also examined, using a linear model with the natural log of the herd's fecal culture prevalence, ln($m_i(\tau)$), or the base-10 log of the average MAP shedding level for the herd, log($f_i(\tau)$), as the response variable; herd was included as a fixed variable. The same three previously defined S6 predictor variables were examined separately.

The sensitivity of S6 testing to detect animals shedding MAP in the herd was calculated using individual FC results as the gold standard. This sensitivity estimate was used to calculate the true US national herd prevalence of MAP from the 68% apparent prevalence estimate produced by the NAHMS survey (USDA-APHIS-VS-CEAH, 2008) and an assumed specificity of 1. True prevalence was estimated with a Bayesian approach (Messam et al., 2008) to the Rogan-Gladen prevalence estimation (Rogan and Gladen, 1978) in WinBUGS1.4 (©2003: Imperial College &MRC, UK), with a run-in of 500 iterations, a sample of 70,000 iterations, and a non-informative Beta(1,1) prior.

*Model Fitting*

All models were fit with the *lmer* function in the *lme4* package (Bates and Maechler, 2010) for R 2.11.1 (R Development Core Team, 2010), which was accessed through the Revolution R Analytics interface (© 2010 Revolution Analytics, Inc.). This function fits subject-specific generalized linear mixed models. All interactions were included, and the models were fit with backward selection using BIC comparison; model hierarchy was maintained. All statistical tests were considered

significant at the 0.05 level, with no adjustment for multiple testing. All proportions were natural log transformed, with 0.05 added to all measures to avoid infinite estimates after transformation; all MAP cfus were log-10 transformed, with 0.5 added to all measures to avoid infinite estimates after transformation. Goodness of fit was determined for logistic models with visual comparison of predicted versus observed values using the plot.logistic.fit.fnc function in the languageR package; for linear models, goodness of fit was determined with visual observation of Q-Q plots of the residuals.

**Results**

*All samples*

A total of 1131 EC results were recorded on the 3 farms during the study period, 133 (11.8%) of which were positive for MAP. Of these samples, 545 (125 or 22.9% positive) had concurrent FC results and 383 (47 or 12.3% positive) were associated with FC results from the previous quarter. The distribution of results across environmental sample type in the 3 herds is shown in Table 3.1. Figure 3.1 shows a time series for the average amount of MAP in FCs and the number and results of ECs in the 3 study herds. There is variation in the number of samples included, as on some

Table 3.1: Distribution of 1131 environmental samples by type for 3 commercial dairy herds in the Northeast US; samples were collected quarterly over a 5 year period (2004-2009) and cultured for *Mycobacterium avium* subsp. *paratuberculosis*

| | Number of sample type (percent positive) | | | |
|---|---|---|---|---|
| **Herd** | **Alley** | **Pen[1]** | **Pit[2]** | **Other[3]** |
| A | 27 (11) | 125 (15) | 16 (13) | 170 (0) |
| B | 94 (10) | 48 (10) | 22 (18) | 43 (5) |
| C | 51 (25) | 260 (18) | 86 (35) | 189 (0) |

[1]Pens containing adult cows
[2]Manure storage areas, including pits and spreaders
[3]All other samples, including feed, water, and calf housing areas

sampling dates a greater number of ECs was collected for reasons associated with other research questions, while on others some of the samples were contaminated and could not be included in the results.

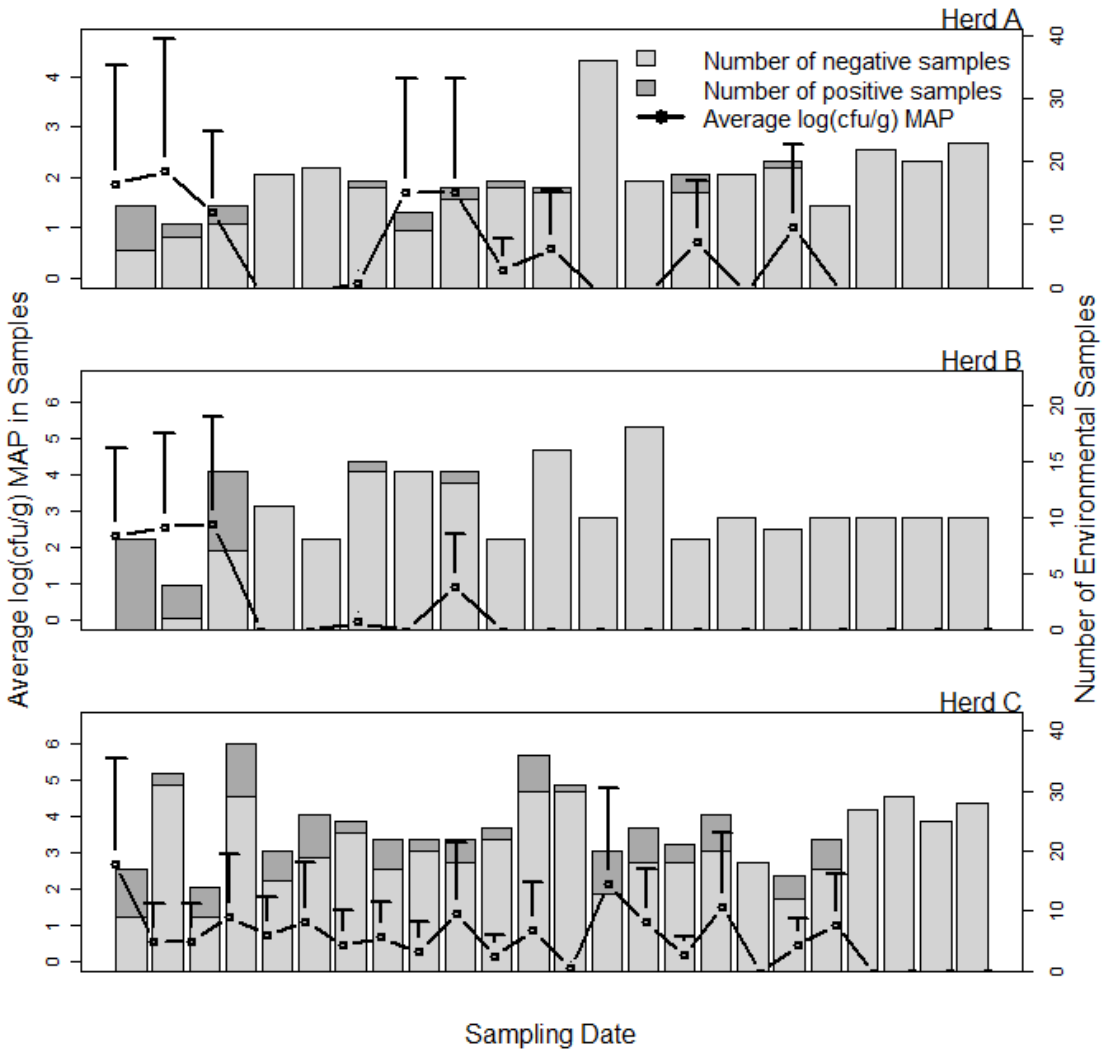



Figure 3.1: Timeline of *Mycobacterium avium* subsp. *paratuberculosis* (MAP) sampling intensity and results for each of 3 commercial US dairy herds between 2004 and 2009. Bars represent the total number of environmental samples collected in the herd at that sampling date, with positive samples represented in dark grey and negative samples in light gray. The base-10 log of the average amount of MAP (cfu/g) in individual fecal samples is represented by the black line, with error bars representing the base-10 log of the standard deviation of MAP (in cfu/g) in individual fecal samples as a rough indicator of the variability in samples.

The models using the full database were found to be poorly fitted, with normality and linearity assumptions failing. Thus, analysis was focused on the separate datasets of pen samples and non-pen samples, which were able to meet statistical assumptions and produce well-fitted models.

There were 250 ECs from adult animal pens with full data for analysis, 42 of which were positive for MAP (16.8%). The average number of high shedders in a pen at any timepoint was 0.27 for herd A (median=0, range 0 to 4), 0 for herd B, and 0.25 for herd C (median=0, range 0 to 1). Results of the logistic regression (Equation 3.1) for pen sample data are shown in Table 3.2; linearity assumptions for explanatory

Table 3.2: Results of logistic regression for the probability of a *Mycobacterium avium* subsp. *paratuberculosis* (MAP)-positive environmental sample culture from an adult cow pen by average MAP shedding level (log(average cfu/g), above), the proportion of animals in the pen shedding MAP, or the number of animals in the pen shedding ≥30 cfu/g MAP in their feces, in 3 commercial US dairy herds, either on the date of sampling or in the previous quarter, between 2004 and 2009. Sampling date is included as a random variable[a].

| Variable | Estimate (OR) | Standard Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| ***Logistic Regression based on Fecal Shedding*** | | | | |
| **(Intercept)** | -1.98 | 0.26 | -7.71 | <0.01 |
| **Average fecal shedding (log cfu/g)** | 0.88 (2.41) | 0.29 | 3.03 | <0.01 |
| ***Logistic Regression based on Fecal Prevalence*** | | | | |
| **(Intercept)** | 0.40 | 0.93 | 0.42 | 0.67 |
| **ln(fecal prevalence[b])** | 0.86 (2.36) | 0.34 | 2.50 | 0.01 |
| ***Logistic Regression based on High-Shedders*** | | | | |
| **(Intercept)** | -2.05 | 0.26 | -7.96 | <0.01 |
| **high shedders[c]** | 0.63 (1.88) | 0.32 | 1.98 | 0. 05 |

[a]Sampling date had a variance of 0.83 for the shedding model, 0.72 for the prevalence model, and 0.68 for the high-shedders model.
[b]expressed as a proportion, natural log (base e)
[c]number of animals shedding ≥30 cfu/g MAP in their feces

variables were met for all models. A 1-log increase in average shedding level or an increase of 0.027 in the prevalence of fecal shedders in a pen resulted in a 2.4-fold increase in the odds of a positive sample, and 1 additional high shedder resulted in a 2-fold increase. Herd effects were not significant for any model.

Median distance between pens was 43m for herd A (range 16 to 108), 39m for herd B (range 36 to 54), and 31m for herd C (range 23 to 48). The linear models based on prevalence and number of high-shedders in the pen did not meet normality assumptions for explanatory variables and random errors and produced poorly-fitted models; the results of the linear model based on average fecal shedding in the pens is shown in Table 3.3. A 2 log unit increase in the average shedding level of animals in a pen was required to raise the contamination of samples in that pen by 1 log. The effect of MAP shedding by cows in other pens was not significantly related to the amount of MAP in a pen's environment. Herd effects were not significant for any model.

There were 542 samples available for analyzing the results of EC from non-pen sources, of which 55 (10.1%) were positive for MAP. Results of the logistic and linear regressions based on whole-herd prediction variables are shown in Figure 3.2 and Tables 3.4 and 3.5, respectively. A 1-log increase in average shedding increased the odds of finding MAP 6-fold and the average amount of MAP by 0.5-log in alleyway or pit samples. Fecal prevalence had a positive relationship with both the probability of finding MAP and the average amount of MAP in environmental samples, but this relationship was strongest in alleyway samples. The number of high-

shedding animals was not significantly related to the odds of finding MAP or the amount of MAP in a sample. Herd effects were not significant for any model.

Table 3.3: Results of a linear regression for *Mycobacterium avium* subsp. *paratuberculosis* (MAP) concentration (log(cfu/g)) in environmental samples from adult cow pens in 3 commercial US dairy herds, either on the date of sampling or in the previous quarter, between 2004 and 2009. Sampling date is included as a random variable[a].

| Variable | Level | Estimate | Std. Error | t value | Pr(>t) | $\chi^2$ (Pr> $\chi^2$) |
|---|---|---|---|---|---|---|
| Intercept | | 0.60 | 0.11 | 5.32 | <0.01 | |
| pen shedding[b], log(cfu/g) | | 0.54 | 0.07 | 8.01 | <0.01 | |
| neighbor shedding[c], log(cfu/g)/m$^2$ | | 35.70 | 38.67 | 0.92 | 0.25 | |
| Pen (calving is base) | 1 | -0.59 | 0.10 | -5.74 | <0.01 | 70.33 (<0.01) |
| | 2 | -0.67 | 0.11 | -6.28 | <0.01 | |
| | 3 | -0.64 | 0.10 | -6.24 | <0.01 | |
| | 4 | -0.86 | 0.11 | -8.03 | <0.01 | |
| | dry | -0.61 | 0.11 | -5.68 | <0.01 | |

[a]Sampling date had a variance of 0.09.
[b]Average shedding level for animals in the pen at the time of sampling.
[c]Sum of average shedding level for animals in other pens divided by squared distance (in m) between the center of the sampled pen and the center of the other pens.

For the distance-corrected linear regression with non-pen samples, pit samples were excluded; 249 samples remained, of which 13 samples from alleyways and 0 samples from other sites were positive for MAP (5.2%). Median distance between non-pen sampling sites and any pen was 78m for herd A (range 19:174), 33m for herd B (range 11:93), and 50m for herd C (range 11:131). Linearity assumptions were not met for these models (results not shown).

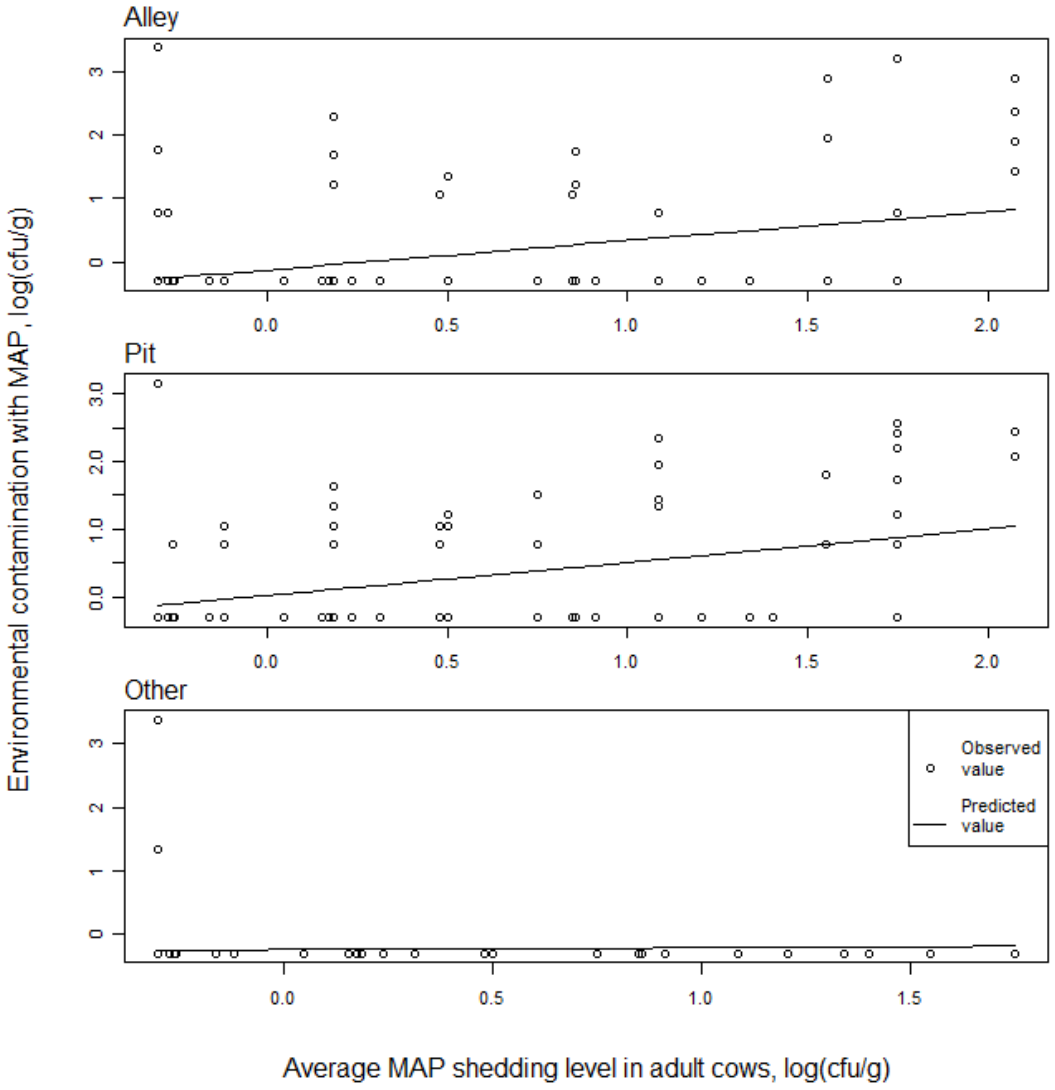



Figure 3.2: Predicted (lines) and observed (symbols) amounts of environmental contamination by *Mycobacterium avium* subsp. *paratuberculosis* (MAP) (in log(cfu/g)) in 3 commercial US dairy herds between 2004 and 2009 based on the sample type (shared alleys; manure pits; all other) and the average fecal sampling of MAP in all adult animals in the herd in the current or previous quarters (in log(cfu/g)). Shedding levels were log-transformed (base-10) after adding 0.5 to all values.

Table 3.4: Results of logistic regression for the probability of a *Mycobacterium avium* subsp. *paratuberculosis* (MAP)-positive environmental sample culture from a source other than adult cow pens by average MAP shedding level in the herd (log(average cfu/g)), the proportion of animals in the herd shedding MAP, or the number of animals in the herd shedding ≥30 cfu/g MAP in their feces, in 3 commercial US dairy herds, either on the date of sampling or in the previous quarter, between 2004 and 2009. Sampling date is included as a random variable[a].

| Variable | | Estimate (OR) | Standard Error | z value | Pr(>\|z\|) | $\chi^2$ (p> $\chi^2$) |
|---|---|---|---|---|---|---|
| ***Logistic Regression based on Fecal Shedding*** | | | | | | |
| **(Intercept)** | | -3.34 | 0.49 | -6.8 | <0.01 | |
| **Average fecal shedding (log cfu/g)** | | 1.78 (5.93) | 0.44 | 4.0 | <0.01 | |
| **Type (Alley is base)** | ***Other*** | -3.83 (0.02) | 0.94 | -4.1 | <0.01 | 78.32 (<0.01) |
| | ***Pit*** | 0.84 (2.32) | 0.44 | 1.9 | 0.06 | |
| ***Logistic Regression based on Fecal Prevalence*** | | | | | | |
| **(Intercept)** | | 8.8 | 3.1 | 2.8 | <0.01 | |
| **ln(prevalence)** | | 4.5 | 1.2 | 3.6 | <0.01 | |
| **Type (Alley is base)** | ***Other*** | -398.8 | 73426 | 0.0 | 1.00 | 21.39 (<0.01) |
| | ***Pit*** | -1.8 | 3.1 | -0.6 | 0.56 | |
| **ln(prevalence): Type** | ***Other*** | -133.4 | 24510 | 0.0 | 1.00 | 15.77 (<0.01) |
| | ***Pit*** | -1.1 | 1.3 | -0.8 | 0.41 | |
| ***Logistic Regression based on High-Shedders*** | | | | | | |
| **(Intercept)** | | -3.68 | 0.76 | -4.8 | <0.01 | |
| **high shedders[b]** | | 0.64 (1.90) | 0.40 | 1.6 | 0.11 | |
| **Type (Alley is base)** | ***Other*** | -3.24 (0.04) | 0.99 | -3.3 | <0.01 | 54.12 (<0.01) |
| | ***Pit*** | 1.28 (3.6) | 0.57 | 2.3 | 0.02 | |

[a]Sampling date had a variance of 2.90 for the shedding model, 3.43 for the prevalence model, and 4.70 for the high-shedders model.
[b]number of animals shedding ≥30 cfu/g MAP in their feces

Table 3.5: Results of a linear regression for *Mycobacterium avium* subsp. *paratuberculosis* (MAP) concentration (log(cfu/g)) in environmental samples from a source other than adult cow pens by average MAP shedding level in the herd (log(average cfu/g)), the proportion of animals in the herd shedding MAP, or the number of animals in the herd shedding ≥30 cfu/g MAP in their feces, in 3 commercial US dairy herds, either on the date of sampling or in the previous quarter, between 2004 and 2009. Sampling date is included as a random variable[a].

| Variable | Level | Estimate | Std. Error | t value | Pr(>t) | $\chi^2$ (Pr> $\chi^2$) |
|---|---|---|---|---|---|---|
| ***Linear Regression based on Fecal Shedding*** | | | | | | |
| **Intercept** | | -0.13 | 0.07 | -2.0 | 0.06 | |
| **average fecal shedding[b], log(cfu/g)** | | 0.46 | 0.09 | 5.4 | <0.01 | |
| **Type** | ***Other*** | -0.11 | 0.06 | -2.0 | 0.06 | 12.06 |
| **(Alley is base)** | ***Pit*** | 0.15 | 0.07 | 2.0 | 0.06 | (<0.01) |
| **shedding:Type** | ***Other*** | -0.43 | 0.08 | -5.3 | <0.01 | 34.69 |
| **(Alley is base)** | ***Pit*** | 0.03 | 0.09 | 0.4 | 0.35 | (<0.01) |
| ***Linear Regression based on Fecal Prevalence*** | | | | | | |
| **Intercept** | | 2.21 | 0.49 | 4.5 | <0.01 | |
| **ln(prevalence)** | | 0.83 | 0.18 | 4.5 | <0.01 | |
| **Type** | ***Other*** | -2.80 | 0.44 | -6.4 | <0.01 | 63.85 |
| **(Alley is base)** | ***Pit*** | -0.02 | 0.48 | -0.04 | 0.39 | (<0.01) |
| **ln(prevalence):Type** | ***Other*** | -0.97 | 0.16 | -6.0 | <0.01 | 49.43 |
| **(Alley is base)** | ***Pit*** | -0.07 | 0.18 | -0.4 | 0.36 | (<0.01) |
| ***Linear Regression based on High Shedders*** | | | | | | |
| **Intercept** | | -0.11 | 0.12 | -1.0 | 0.23 | |
| **number of high shedders** | | 0.14 | 0.08 | 1.7 | 0.10 | |
| **Type** | ***Other*** | -0.05 | 0.10 | -0.5 | 0.34 | 5.12 |
| **(Alley is base)** | ***Pit*** | 0.21 | 0.12 | 1.7 | 0.10 | (0.08) |
| **high shedders:Type** | ***Other*** | -0.17 | 0.07 | -2.5 | 0.03 | 12.13 |
| **(Alley is base)** | | | | | | (<0.01) |
| | ***Pit*** | 0.04 | 0.08 | 0.5 | 0.34 | |

[a]Sampling date had a variance of 0.13 for the shedding model, 0.18 for the prevalence model, and 0.19 for the high-shedder model.
[b]Average shedding level for animals in the herd at the time of sampling.

*VBJDCP samples*

Of the 483 collected environmental samples meeting the criteria for the S6, 92 were positive for MAP (19.0%). There were 50 sampling dates with both FC results and at least 6 S6 samples from the same or previous quarter.

The results of the logistic model for VBJDCP results showed that none of the S6 results (dichotomous, count, or average contamination) were significant predictors of MAP presence in the herd (p=0.998 for each variable, data not shown). Likewise, none of the S6 predictor variables were significantly related to the FC prevalence in the herd (p=0.4 for each variable, data not shown), nor was herd a significant variable in any of the models. However, all S6 predictor variables were positively correlated with the average fecal shedding in the herd (Table 3.6).

The S6 results correctly identified as positive 19 of 47 dates (40%) in which animals were shedding MAP (and correctly identified as negative 3 of 3 dates with no positive MAP FC samples). This indicates a relative sensitivity of 0.40 (95% CI: 0.26 to 0.54) compared to whole herd FC, and supported our assumption of a specificity of 1. The NAHMS environmental survey, using similar sampling methods to our study, found a MAP prevalence of 0.681 among dairy operations. Assuming the sensitivity of environmental sampling observed in our study, the true herd prevalence of MAP would therefore be 0.977 (95% CI: 0.921 to 0.983).

Table 3.6: Results of linear regression for average fecal shedding, log(cfu/g), of *Mycobacterium avium* subsp. *paratuberculosis* (MAP) in individual animals in commercial US dairy herds between 2004 and 2009 by results of 6 VBJDCP-standard environmental samples (the S6), by the dichotomous results (S6 result, top), the total number of positive samples (Positive count, middle) or the average MAP concentration in the samples (mean log(cfu/g), bottom)

| Variable | Level | Estimate | Std. Error | t value | Pr(>\|t\|) | Residual Deviance |
|---|---|---|---|---|---|---|
| | | | ***Dichotomous Model*** | | | |
| **(Intercept)** | | 0.79 | 0.14 | 4.84 | <0.01 | 14.02 |
| **S6 result** | | 0.48 | 0.20 | 2.42 | 0.02 | |
| **Farm** | ***B*** | -0.83 | 0.19 | -4.39 | <0.01 | |
| **(A is base)** | ***C*** | -0.47 | 0.22 | -2.18 | 0.03 | |
| | | | ***Count Model*** | | | |
| **(Intercept)** | | 0.74 | 0.13 | 5.73 | <0.01 | 12.69 |
| **Positive count** | | 0.17 | 0.05 | 3.38 | <0.01 | |
| **Farm** | ***B*** | -0.92 | 0.18 | -5.21 | <0.01 | |
| **(A is base)** | ***C*** | -0.46 | 0.19 | -2.37 | 0.02 | |
| | | | ***Concentration Model*** | | | |
| **(Intercept)** | | 0.76 | 0.13 | 5.89 | <0.01 | 12.71 |
| **log(mean cfu)** [a] | | 0.27 | 0.08 | 3.36 | <0.01 | |
| **Farm** | ***B*** | -0.83 | 0.18 | -4.68 | <0.01 | |
| **(A is base)** | ***C*** | -0.37 | 0.19 | -2.01 | 0.05 | |

[a]Log (base-10) transformation

***Discussion***

Our study showed that environmental contamination with MAP was significantly correlated with MAP shedding levels in individual animals. Other studies have considered the relationship between environmental contamination and fecal shedding in dairy cattle, with similar findings. Contamination of bedding with *Klebsiella pneumoniae* has been associated with fecal shedding in animals using the bedding, but only a subset of animals were sampled and there was insufficient

variation in *Klebsiella* concentration in bedding to allow further analysis (Munoz et al., 2006). Environmental sample positivity in an endemically infected herd has appeared to correspond temporally with active Salmonella shedding (Wray et al., 1989). The current study shows that average shedding level, as well as shedding prevalence, is important in understanding the amount of environmental MAP contamination via feces. Pillars et al. (2009) likewise found a relationship between prevalence and environmental MAP contamination, but did not consider either the amount of fecal shedding or the amount of environmental contamination. The results presented here usefully extend their findings; the presence of one or two animals with extremely high shedding rates (super shedders) could bias a model based only on prevalence.

This study confirmed that the environment of known MAP positive dairy farms can be contaminated by animals shedding MAP, though the proportion of positive samples was lower than in previous studies (Lombard et al., 2006; Pillars et al., 2009); the difference is likely a combination of herd prevalence (which was low for 2 of the 3 herds in this study) and sample types (which were more varied in the current study than in previous studies). Certain environments on the farm were more likely to be contaminated with MAP and had higher average contamination levels. These high risk areas included manure storage areas and shared alleyways, parts of the farm in which the manure of adult cows are mixed. High MAP concentration in fecal cultures increased the amount of MAP found in manure storage areas, which seems a logical and expected result; previous studies have found that MAP can survive in manure storage areas for >200 days (Jorgensen, 1977; Lovell et al., 1944). Figure 3.2

shows that there were a large number of negative environmental samples even when high contemporary fecal shedding levels were measured; these negative environmental samples were primarily from areas labeled 'Other' (Table 3.1), which were not frequented by adult cattle or in contact with their manure and which were often not sampled in previous studies of environmental contamination. However, negative EC samples from alleys and manure collection areas were also observed in some cases with high contemporary average FC shedding levels (Figure 3.2).

Our data lacked quarterly fecal sampling, limiting our analysis to the effects of either concurrent samples or samples from the previous quarters. A recent longitudinal study found that the number of MAP-positive EC samples increased with increasing FC prevalence, but the converse was not consistently true, indicating that some environmental contamination may remain despite successful reduction of incidence (Pillars et al., 2009). Therefore, we chose to include fecal sampling results from the previous quarter as a proxy for the current shedding level. As these herds were not yet provided with FC results from the previous quarter during the quarterly EC sampling, little or no test-based culling would be present to bias this assumption.

We were interested in the effect of proximity of shedding animals on MAP contamination in farm environments, and we found that samples taken from freestall pens, in which adult animals directly shared an environment, were significantly related to the results of individual FC inside those pens. The results of pen sampling were correlated with the concurrent presence of high-shedding animals. As these animals have lower milk production (Nielsen et al., 2008; Smith et al., 2009) and longer calving intervals (Smith et al., 2010), it may be economically desirable to detect their

presence and remove them. Additionally, high shedding animals pose an important risk for other MAP susceptible animals in the herd (Lu et al., 2010; Pradhan et al., 2011). Our study observed a correlation between the probability of finding MAP in adult cow pens and the presence of high shedders, suggesting that regular sampling of adult cow pens may be a good method for detecting the presence of highly infectious animals, in agreement with Aly et al. (2009).

We were also interested in the ability of shedding animals in the adult cow pens to contaminate other environments on the farm. Specifically, we wanted to identify a link between EC results for non-pen/non-pit samples and the distance-corrected shedding level in the herd. Such a link could not be identified in the current data, as statistical assumptions were not met, although a relationship was observed between all non-pen samples and the herd-level results of FC; there were very few positive non-pen/non-pit environmental samples, so the power of the distance-corrected analysis was low. We also were limited to straight-line distance analyses, ignoring walls and fences, which does not necessarily reflect traffic patterns and other contamination methods. In addition, on one farm (Herd C), alleyways frequently passed through adult cow pens, making distance calculations inappropriate. However, the low number of positive samples suggests that the level of MAP contamination in feed bunks, water sources, and other such locations was negligible; this in itself suggests that the optimal use of hygiene program resources is to focus on MAP contamination of maternity pens and other known transmission methods.

Our study showed that a single S6 sampling had a relative sensitivity of only 0.40 compared to whole-herd FC; there were only 3 sampling dates with no positive

fecal cultures, and the herds were known to be endemically infected, so specificity calculations would be inappropriate. These results agree with the simulations of Tavornpanich et al. (2008), which estimated the herd sensitivity of this testing method to be 0.38 with a within-herd prevalence of 0.05; the 3 herds in this study had a prevalence of 0.05 for the majority of sampling dates. This has important implications for the interpretation of results from this sampling method; a negative result by this method does not necessarily mean that the sampled herd is MAP-free. The relative sensitivity of environmental culture was calculated based on the imperfect FC test, which is known to have low sensitivity in low-shedding animals (Collins et al., 2006); this would lead to a still lower true sensitivity for the VBJDCP protocol. The S6 samples were unable to significantly predict the presence of MAP in the feces of adult animals on the farm, nor the FC prevalence of MAP in the herd, although both of these could have been sensitive to the low number of sampling dates with no positive FC and the overall low prevalence of MAP in the herds. However, the S6 samples, especially their average MAP cfu's, were able to predict well the average fecal shedding of the herd. This would indicate that the S6 method is sensitive to the shedding level of animals within a herd, and may not be able to detect herds in which fecal shedding is low; this was noted by Pillars et al. (2009), who could not culture MAP from environmental samples taken from herds with a prevalence of <0.02, and by Raizman et al. (2004), who found that infected herds with 2 negative environmental cultures had ≤0.04 prevalence by pooled random sampling. As this category includes newly-infected herds, in which the majority of animals would be latently infected or

low-shedding, S6 results may not be sufficient to detect herds with new MAP infections.

This study shows that the presence of MAP in the environment of a farm or pen is correlated with fecal shedding in the cows; however, with our diagnostic methods, MAP may be absent from environmental samples despite high levels of fecal shedding in the cows.

***Acknowledgements***

The USDA (Cooperative State Research, Education and Extension Service, Washington, D.C.) Award Number 2008-35204-04627 provided funding for this study, as did the USDA-Agricultural Research Service (Agreements. 58-1265-3-155, 58-1265-3-156, 58-1265-3-158, and 58-1265-4-020) for the Regional Dairy Quality Management Alliance (RDQMA) and the Johne's Disease Integrated Program (JDIP, USDA contract 45105). We are grateful to the comments of the reviewers and the associate editor, which have greatly improved our analysis.

# REFERENCES

Aly, S. S., R. J. Anderson, R. H. Whitlock, T. L. Fyock, S. McAdams, J. M. Adaska, J. Jiang, and I. A. Gardner. 2009. Reliability of environmental sampling to quantify Mycobacterium avium subspecies paratuberculosis on California free-stall dairies. J. Dairy Sci. 92:3634-3642.

Bates, D., and M. Maechler. 2010. lme4: Linear mixed-effects models using S4 classes.

Benedictus, A., R. M. Mitchell, M. Linde-Widmann, R. W. Sweeney, T. L. Fyock, Y. H. Schukken, and R. H. Whitlock. 2008. Calculation of transmission parameters of Mycobacterium avium subspecies paratuberculosis infections in a dairy herd going through a control program. Prev. Vet. Med. 83(3-4):215-227.

Berghaus, R. D., T. B. Farver, R. J. Anderson, C. C. Jaravata, and I. A. Gardner. 2006. Environmental Sampling for Detection of Mycobacterium avium ssp. paratuberculosis on Large California Dairies. J. Dairy Sci. 89(3):963-970.

Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, Jr., and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229(12):1912-1919.

Jorgensen, J. B. 1977. Survival of Mycobacterium paratuberculosis in slurry. Nord. Vet. Med. 29(6):267-270.

Lombard, J. E., B. A. Wagner, R. L. Smith, B. J. McCluskey, B. N. Harris, J. B. Payeur, F. B. Garry, and M. D. Salman. 2006. Evaluation of Environmental Sampling and Culture to Determine Mycobacterium avium subspecies paratuberculosis Distribution and Herd Infection Status on US Dairy Operations. J. Dairy Sci. 89(11):4163-4171.

Lovell, R., M. Levi, and J. Francis. 1944. Studies on the survival of Johne's bacilli. J. Comp. Pathol. 54:120-129.

Lu, Z., Y. H. Schukken, R. L. Smith, and Y. T. Grohn. 2010. Stochastic simulations of a multi-group compartmental model for Johne's disease on US dairy herds with test-based culling intervention. J. Theo. Bio. 264(4):1190-1201.

Messam, L. L. M., A. J. Branscum, M. T. Collins, and I. A. Gardner. 2008. Frequentist and Bayesian approaches to prevalence estimation using examples from Johne's disease. Anim. Health Res. Rev. 9(1):1-23.

Munoz, M. A., C. Ahlström, B. J. Rauch, and R. N. Zadoks. 2006. Fecal shedding of Klebsiella pneumoniae by dairy cows. J. Dairy Sci. 89:3425-3430.

Nielsen, S. S., M. A. Krogh, and C. Enevoldsen. 2008. Time to the occurence of a decline in milk production in cows with various paratuberculosis antibody profiles. J. Dairy Sci. 92:149-155.

Nielsen, S. S., and N. Toft. 2009. A review of prevalences of paratuberculosis in farmed animals in Europe. Prev. Vet. Med. 88:1-14.

Pillars, R. B., D. L. Grooms, and J. B. Kaneene. 2009. Longitudinal study of the distribution of Mycobacterium avium subsp. paratuberculosis in the environment of dairy herds in the Michigan Johne's disease control demonstration herd project. Can. Vet. J. 50:1039-1046.

Pradhan, A., J. S. Van Kessel, J. S. Karns, D. R. Wolfgang, E. Hovingh, K. A. Nelen, J. M. Smith, R. H. Whitlock, T. L. Fyock, S. Ladely, P. J. Fedorka-Cray, and Y. H. Schukken. 2009. Dynamics of endemic infectious diseases of animal and human importance on three dairy herds in the Northeastern US. J. Dairy Sci. 92:1811-1825.

Pradhan, A. K., R. M. Mitchell, A. J. Kramer, M. J. Zurakowski, T. L. Fyock, R. H. Whitlock, J. M. Smith, E. Hovingh, J. S. Van Kessel, J. S. Karns, and Y. H. Schukken. 2011. Molecular epidemiology of Mycobacterium avium subsp. paratuberculosis in a longitudinal study of three dairy herds. J. Clin. Microbiol. 49(3):893-901.

R Development Core Team. 2010. R: A language and environment for statistical computing. in R Foundation for Statistical Computing, Vienna, Austria.

Raizman, E. A., S. J. Wells, S. M. Godden, R. F. Bey, M. J. Oakes, D. C. Bentley, and K. E. Olsen. 2004. The distribution of Mycobacterium avium ssp. paratuberculosis in the environment surrounding Minnesota dairy farms. J. Dairy Sci. 87(9):2959-2966.

Rogan, W. J., and B. Gladen. 1978. Estimating prevalence from the results of a screening test. Am. J. Epidemiol. 107(1):71-76.

Smith, R. L., Y. T. Grohn, A. K. Pradhan, R. H. Whitlock, J. S. Van Kessel, J. M. Smith, D. R. Wolfgang, and Y. H. Schukken. 2009. A longitudinal study for the impact of Johne's Disease status on milk production in individual cows. J. Dairy Sci. 92(6):2653-2661.

Smith, R. L., R. L. Strawderman, Y. H. Schukken, S. J. Wells, A. K. Pradhan, L. A. Espejo, R. H. Whitlock, J. S. Van Kessel, J. M. Smith, D. R. Wolfgang, and Y. T. Grohn. 2010. The effect of Johne's disease status on reproduction and culling in dairy cattle. J. Dairy Sci. 93:3513-3524.

Tavornpanich, S., C. A. Munoz-Zanzi, S. J. Wells, E. A. Raizman, T. E. Carpenter, W. O. Johnson, and I. A. Gardner. 2008. Simulation model for evaluation of testing strategies for detection of paratuberculosis in Midwestern US dairy herds. Prev. Vet. Med. 83:65-82.

USDA-APHIS-VS-CEAH. 2008. Johne's Disease on U.S. Dairies, 1991-2007. in Center for Epidemiology and Animal Health, Fort Collins, CO.

USDA:APHIS:VS. 2008. NAHMS 2007: Johne's Disease on U.S. Dairy Operations. Pages 1-4 in USDA:APHIS:VS:NAHMS, Fort Collins, CO.

USDA:APHIS:VS. 2010. Uniform Program Standards for the Voluntary Bovine Johne's Disease Control Program. in USDA-Animal and Plant Health Inspection Service-Veterinary Services, Washington,D.C.

Wray, C., Q. C. Wadsworth, D. W. Richards, and J. H. Morgan. 1989. A three-year study of Salmonella dublin infection in a closed dairy herd. Vet. Rec. 124:532-535.

# CHAPTER 4

# ESTIMATING TRANSMISSION RATE PARAMETERS FOR ENDEMIC CHRONIC PATHOGENS USING A REVERSIBLE-JUMP MARKOV CHAIN MONTE CARLO MODEL: PARATUBERCULOSIS IN DAIRY CATTLE


## *Abstract*

A Bayesian model was developed to estimate the transmission rates from two heterogeneous classes infected with an endemic chronic pathogen of dairy cattle, *Mycobacterium avium* subsp. *paratuberculosis* (MAP), using reversible-jump Markov Chain Monte Carlo (rjMCMC) methods. Unobserved state transition times were treated as nuisance parameters. As MAP diagnostics have poor sensitivity, infection status of test-negative animals was estimated with a hidden Markov model. The model was validated with simulated data designed to mimic longitudinal field data from a commercial dairy herd. The hidden Markov model implemented with the rjMCMC method was shown to be incapable of accurately estimating transmission rates for an endemically infected population due to imperfect diagnostics; the Bayesian model implemented with the MCMC method on a full simulated dataset (all transition times observed) was also incapable of accurately estimating transmission rates due to differences in model assumptions. This model appears to be an inappropriate method for estimating the true transmission rate of MAP in commercial dairy herds.


## *Introduction*

Endemic chronic infectious diseases can be difficult to characterize mathematically, especially when imperfect diagnostic tests result in a large number of

undiagnosed cases. In these circumstances, the technique most capable of producing stable, accurate results is the rjMCMC model. These models are useful in situations in which multiple inter-related parameters must be estimated simultaneously using sparse data (Becker and Britton, 1999; Streftaris and Gibson, 2004b). In addition, rjMCMC models can be used to convert discrete data, such as annual test results, into unobserved continuous variables, such as disease status (Green, 2003; Neal and Roberts, 2004; O'Neill, 2002; O'Neill and Becker, 2001; O'Neill and Roberts, 1999), and can correct for low diagnostic sensitivity by allowing for a test-negative individual to be classified as infected by the process of reversible jumps (Auranen et al., 2000; Forrester et al., 2007).

While rjMCMC techniques have been developed to estimate mathematical model parameters for many infectious diseases with substantial latent periods (Cauchemez and Ferguson, 2008; Lekone and Finkenstädt, 2006; Streftaris and Gibson, 2004a) or imperfect diagnostics (Forrester et al., 2007; Glass et al., 2007), these techniques have only been applied to epidemic situations. In these scenarios, the change in the infectious pressure over time due to the change in the proportion of infected allows for close pinpointing of transmission rates with minimal data. In contrast, endemic diseases frequently show relatively small changes in prevalence, leading to a lack of variation in infectious pressure. As the few observed changes tend to unfold over long timeframes, high levels of correlation may be observed between the estimated transmission rates of heterogeneously infectious groups, or between the transmission rate and the true status of test-negative individuals. In order to study

endemic diseases, especially with imperfect diagnostics, longitudinal studies with repeated testing are likely to be necessary.

In human medicine, these longitudinal studies may be economically or politically difficult; in these circumstances, livestock may serve as a model to first develop the necessary techniques (Lanzas et al., 2010). One livestock example of an endemic, chronic infectious disease is Mycobacterium avium subsp. paratuberculosis (MAP), the causative agent of Johne's disease. Ruminants exposed to MAP can develop chronic progressive intestinal infections leading to decreased milk production (Smith et al., 2009), wasting, and early culling (Smith et al., 2010). A long latent period (Mitchell et al., 2008) and imperfect diagnostic tests (Collins et al., 2006) have made detection of MAP infection difficult, requiring long follow-up times and repeated testing. Estimates from experimental trials have been made of the duration of latency (Rankin, 1961) and the rate of disease progression (van Schaik et al., 2003). However, experimental trials fail to reproduce field transmission conditions, as cost issues prevent these trials from recreating all levels and types of infectious exposure. Previous attempts to estimate parameters for MAP transmission models have relied on large datasets, but have been limited by the poor sensitivity of diagnostic tests (Collins et al., 2006; Whitlock et al., 2000; Whitlock et al., 2007). A model describing MAP transmission in dairy herds shows the importance of longitudinal studies; infectious pressure is dependent on the presence of animals shedding MAP transiently or intermittently, many of which will not be readily detectable (Mitchell et al., 2005; Mitchell et al., 2008). Analyzing field data can be complicated, given the number of

confounding factors and the knowledge gaps caused by imperfect diagnostics and discrete data from continuous systems.

Models for MAP transmission indicate that several classes of infectious animals (transiently shedding, low-shedding, and high-shedding) may be important in infection (Lu et al., 2008; van Roermund et al., 2007; Whitlock et al., 2005); therefore, multiple transmission parameters must be estimated. In dairy herds, the exact source of a MAP infection is generally unknown and must be considered probabilistically. Infection with MAP generally occurs within the first year of age (Benedictus et al., 2008), but the length of the latent period is such that several years are required for antemortem observation of the infection status of calves exposed to MAP. As a result, field trials of MAP must cover long periods of time and analysis must include total animal histories. The data available from commercial dairy herds tend to be the birth date and culling/death date of all animals, the dates at which animals are tested, and the test results.

The objective of this study is to estimate transmission parameters for dynamic mathematical models of MAP commercial dairy herds using an rjMCMC model for the interpretation of longitudinal field trials for endemic chronic infections.

***Materials and Methods***

A model has been previously developed (Lu et al., 2008; Mitchell et al., 2008) to describe the dynamics of MAP in dairy herds; this model was simplified for the purposes of this study. Figure 4.1 shows a flow chart describing the simplified model, with all variables described in Table 4.1. Susceptible calves ($X_1$) may become resistant to infection at 1 year of age ($X_2$), or they may be infected at rate $\lambda$. Infected calves ($Tr$)

transiently shed low levels of MAP until entering latency at rate $\varphi$; latent animals ($H$) become low-shedding animals ($Y_1$) at rate $\sigma$; low-shedding animals become high-shedding animals ($Y_2$) at rate $\nu$. High-shedding animals may be culled for clinical signs at rate $\alpha$; otherwise, all animals over 1 year of age are subject to a constant death/culling rate of $\mu$. The force of infection for transmission, $\lambda$, at time $t$ is

**Equation 4.1**

$$\lambda(t) = \beta_1(Y_1(t) + Tr(t)) + \beta_2 Y_2(t),$$

where $\beta_1$ is the transmission parameter for low-shedding animals and $\beta_2$ is the transmission parameter for high-shedding animals.

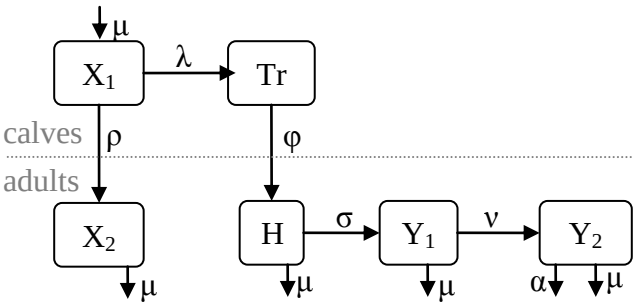



Figure 4.1: Flow chart of the mathematical model for MAP in a dairy herd used to define the rjMCMC model.

Table 4.1: Variables used in the model of paratuberculosis spread in a dairy herd diagrammed in Figure 4.1

| Variable | Description |
|---|---|
| $\mathbf{X_1}$ | susceptible calves |
| $\mathbf{X_2}$ | resistant adults |
| **Tr** | transiently shedding infected heifers |
| **H** | latently infected animals |
| $\mathbf{Y_1}$ | low shedding infected adults |
| $\mathbf{Y_2}$ | high shedding infected adults |

This study will present a Bayesian model designed to include data from all animals in the herd during the study period and will estimate the disease transmission

parameters ($\beta_1$ and $\beta_2$). All calendar times are measured in months since the first whole-herd test.

***Model Specification***

This is a Bayesian hierarchical model for the transmission parameter $\beta_1$, which is the base contribution of MAP-shedding animals to the infection rate; it is assumed that the contribution of high-shedding animals is 4 times the contribution of low-shedding and transiently-shedding animals. The compartmental variables ($\mathbf{X_1}$, **I, Tr,** $\mathbf{Y_1}$, $\mathbf{Y_2}$), henceforward referred to as $\boldsymbol{\pi}$, are functions of the calendar times ($\mathbf{t_b}$, $\mathbf{t_I}$, $\mathbf{t_E}$, $\mathbf{t_{Y1}}$, $\mathbf{t_{Y2}}$, $\mathbf{t_d}$) at which animals enter and leave the compartments ($X_1$, Tr, H, $Y_1$, and $Y_2$, respectively). Of these times, $\mathbf{t_b}$ and $\mathbf{t_d}$, are observed, as are testing result times $\mathbf{t_{t1}}$ and $\mathbf{t_{t2}}$, and we will refer to these observed times as $\boldsymbol{\tau}$; the remaining times ($\mathbf{t_I}$, $\mathbf{t_E}$, $\mathbf{t_{Y1}}$, $\mathbf{t_{Y2}}$), henceforward referred to as $\boldsymbol{\gamma}$, are unobserved. The joint density of the parameter, the observations, and the unobserved variables was

**Equation 4.2**

$$P(\boldsymbol{\tau}, \boldsymbol{\gamma}, \beta_1) = P(\boldsymbol{\tau}|\boldsymbol{\gamma}, \beta_1)P(\boldsymbol{\gamma}|\beta_1)P(\beta_1).$$

***As in Auranen et al. (2000)*** the components of the right-hand side of the equation are referred to as the observation level, the transmission level, and the prior level, respectively. The structure of each level is explicitly stated below.

*Observation level*

The role of this level is to connect the unobserved (augmented) data, $\boldsymbol{\gamma}$, with the observed data, $\boldsymbol{\tau}$. This level is presented at the level of a single animal, *i*. For the sake of clarity, we have omitted the indicator *i*. The joint density of an animal's observed data is expressed as

**Equation 4.3**

$$P(\boldsymbol{\tau}|\boldsymbol{\gamma}, \beta_1) = P(t_{t2}|t_{Y2})P(t_{t1}, t_d|t_{Y1}, t_E)P(t_b|t_I, \beta_1, \boldsymbol{\pi})$$

where the components of the right-hand side of the equation refer to high-shedding, low-shedding, and infection, respectively. This necessarily assumes that $t_{t2}$ is independent of $t_b$, and that both are independent of $t_{t1}$ and $t_d$. This assumption is reasonable as the time of birth is not related to the probability of testing positive given infection and with our previous assumption that all high-shedding animals will have high-positive test results regardless of time of death/culling.

The distribution of an animal's high-positive result is a discrete uniform distribution,

**Equation 4.4**

$$P(t_{t2}|t_{Y2}) = I\{t_{t2} - 12 \leq t_{Y2} \leq t_{t2}\}\left[\frac{1}{12}\right]$$

where I{} denotes the indicator function. This distribution assumed that all high shedding animals had at least one high-positive test result, and that there is annual testing and no false-negative test results.

The distribution of an animal's low-positive test result is the joint distribution of the positive test results given shedding onset, $P(t_{t1},t_d|t_{Y1})$, and the truncated exponential distribution for onset of latency, which is dependent on the experimentally-determined parameter φ:

**Equation 4.5**

$$P(t_{t1}, t_d|t_{Y1}, t_E) = \begin{cases} P(t_{t1}, t_d|t_{Y1})\dfrac{\left(e^{-[tE-tI]/\varphi}\right)/\varphi}{1 - e^{-[tY1-tI]/\varphi}} & if \quad t_{Y1} > 0 \\ \dfrac{\left(e^{-[tE-tI]/\varphi}\right)/\varphi}{1 - e^{-[td-tI]/\varphi}} & else \end{cases}$$

where the distribution of positive test results given shedding onset is the weighted sum of 2 discrete uniform distributions,

**Equation 4.6**

$$P(t_{t1}, t_d | t_{Y1}) = I\{t_{Y1} < t_{t1} < t_d\} * \left[\frac{Se}{min(t_{t1} - t_E, 12)} + I\{t_{t1} > t_E + 12\}\frac{1 - Se}{min(t_{t1} - t_E - 12, 12)}\right]$$

The weighting factor, Se, is the sensitivity of a fecal culture test to low-shedding animals and 1-Se, by extension, is the probability of a single false negative fecal culture test; annual testing is assumed.

The distribution of an animal's birth date with respect to infection is given as

**Equation 4.7**

$$P(t_b | t_I, \beta_1, \boldsymbol{\pi}) = I\{t_b \leq t_I \leq t_b + 12\}\frac{1}{12} + I\{t_I = 0\}\left[\prod_{t=t_b}^{t_b+12} 1 - \lambda(t)\right]$$

with rate of infection λ, as described in Equation 4.1.

*Transmission level*

The distribution of the compartmental variables, which are direct functions of the unobserved data, is assumed to be Poisson distributed,

**Equation 4.8**

$$P(\boldsymbol{\pi} | \beta_1) = \prod_t \prod e^{-X_1(t)\lambda(t)} \frac{[X_1(t)\lambda(t)]^{I(t)}}{I(t)!}$$

where incidence I(t) is the count of infections at time t, $I(t) = sum[which(t_I = t)]$, and $X_1(t)$ and $\lambda(t)$ are as previously defined.

*Prior level*

The prior for the parameter $\beta_1$ was chosen to be uniform on the range [0,n] where n approaches infinity.

***MCMC Sampling***

We used a Markov chain Monte Carlo for estimation, constructed so the stationary distribution is the posterior distribution of the parameters and augmented data given the observed data, denoted $P(\gamma,\beta_1|\tau)$. For each test-positive animal and a proportion $P_{FN}$ of the test-negative animals, $\gamma(i)$ was drawn from the distributions described in the observation level section above, Equations 4.3-4.7. The chain was started with these augmented data and the initial seed was drawn from the uniform distribution U[0.0001,0.005]. We performed random-walk Metropolis sampling (Gilks et al., 1996), which should allow convergence to the desired posterior (Roberts and Tweedie, 1996). At each iteration, the model parameter $\beta_1$ was resampled; if the current value was b, a new value b* was generated so that $b^*=p+\delta u$ and u was drawn from the uniform distribution U[-0.5,0.5] with $\delta$ selected so as to provide sufficient mixing. At every fourth iteration, either the unobserved data $\gamma$ was resampled (with probability 0.5) or a Metropolis sampling was performed in which a test-negative animal was randomly selected and a new $\gamma(i)$ was drawn assuming that $I\{t_I>0\}=1-I\{t_I^*>0\}$.

We performed 50,000 iterations for each run of the MCMC algorithm and the first 1300 were discarded as the burn-in period as recommended by the Raftery diagnostic test (Raftery and Lewis, 1992). The remaining output was recorded to constitute a sample from the posterior distribution. The convergence of the MCMC was tested with the Gelman-Rubin criterion (Gelman and Rubin, 1992): 2 chains were

run with different initial seeds and a Gelman-Rubin Information criterion (GRIC) value under 1.1 was a sign for convergence.

***Validation***

We validated the model using data simulated by an individual-based model using a Gillespie algorithm (Appendix 1). Data was provided from this model for 2 values of $\beta_1$, 0.001 and 0.002, which produced noticeably different predicted prevalences. As this model output included the true transition times, $\boldsymbol{\tau}$, and therefore the compartmental variables, $\boldsymbol{\pi}$, it was possible to model both the reversible-jump model, $P(\boldsymbol{\tau},\boldsymbol{\gamma},\beta)$, and the full-data model,

**Equation 4.9**

$$P(\boldsymbol{\pi}, \beta_1) = P(\boldsymbol{\pi}|\beta_1)P(\beta_1)$$

where $P(\boldsymbol{\pi}|\beta_1)$ is described in Equation 4.8.

As animals that have not been tested for MAP infection do not add information, only animals added to the herd at least 2 years prior to the end of the simulation were included in the analysis. Additionally, only animals added to the herd at least 25 years after the beginning of the simulation were included, so as to include only data from herds in an endemic steady-state. Deviation from the true value of $\beta_1$ was calculated by subtracting the true value from the predicted values in the converged chain. Proportional deviation from the true value was calculated by dividing the deviation by the true value.

***Results***

The posterior density functions of the deviation and proportional deviation, predicted from simulated data, are shown in Figure 4.2. The full data chains have reached convergence

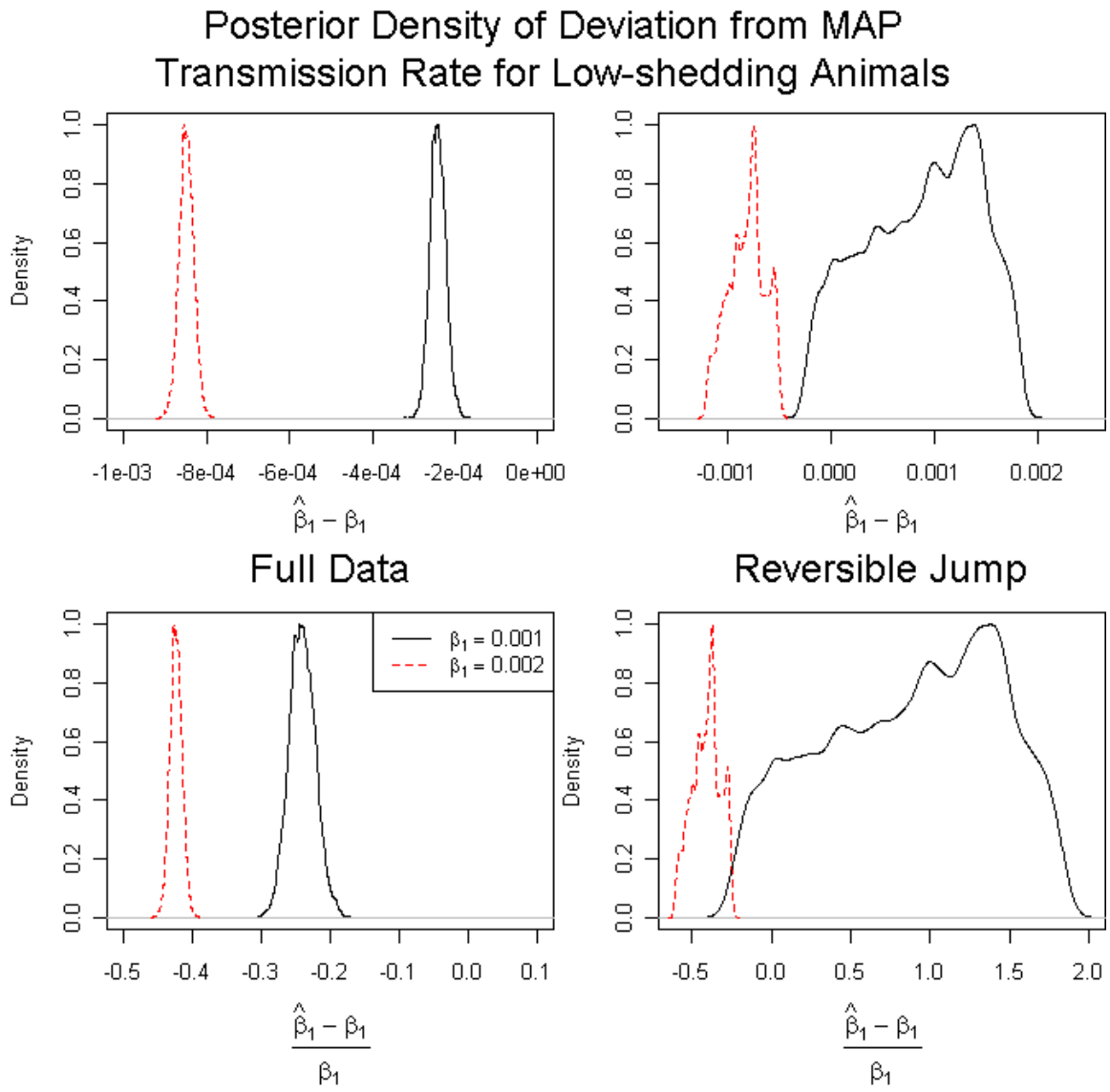



Figure 4.2: Posterior density of deviation (top) and proportional deviation (bottom) from true MAP transmission rate ($\beta_1$=0.001, solid black line; $\beta_1$=0.002, dashed red line) for a full data model (left, Equation 4.9) and a reversible jump model (right, Equation 4.2).

(GRIC < 1.001) but the reversible jump chains have not (GRIC = 1.94 for $\beta 1 = 0.001$, GRIC = 4.12 for $\beta 2 = 0.002$). The recommended chain length for the full data model, based on the Raftery diagnostic test, was more than 83,000 iterations; for the reversible jump model, the recommended chain length was more than 407,000 iterations. As an average iteration of the reversible jump model requires 5 seconds, the time required to reach this length for a single chain is 24 days.

As seen in Figure 4.2, neither model (full data or reversible jump) accurately predicts the true transmission rate. The full data model is able to converge to a stable distribution consistently lower than the true value, although the amount of this deviation is dependent on the true transmission rate. In contrast, the reversible jump model results in poor, slow convergence, with the high transmission rate underestimated and the low transmission rate overestimated.

***Discussion***

The proposed model for estimating MAP transmission rates in dairy cattle clearly is insufficient, even for simulated data. The failure of the full data model is most likely due to some differences between the simulation model and the assumptions of the Bayesian model. The primary difference in assumptions is that of stochastic aging as opposed to deterministic aging: in the simulation model, animals are susceptible for 1 year on average, but may become resistant earlier or later (depending on an exponential distribution with rate = 1 year). In contrast, the Bayesian model assumes that all susceptible animals become resistant at 1 year of age, which could misestimate the number of susceptible animals. This is most important when the transmission rate is high and most animals are infected, as the animals

leaving susceptibility early are more protected from infection; the result is an underestimation of the transmission rate, as observed. It would be interesting to test the Bayesian model using data simulated under the same assumptions, but the time required for that is beyond the scope of this work at this time.

The failure of the reversible jump model is more easily explained: there is simply too little information provided by antemortem testing for MAP. If postmortem testing is added to the model, with an ability to detect latently infected animals, the model converges more quickly, but the resulting posterior distribution is biased in the same way (results not shown). Lacking information on the exact times at which animals become infected and begin shedding, the model is unable to accurately estimate the transmission rate.

As the transmission model-based Bayesian methodology failed to estimate the transmission rate, other estimation methods should be considered for this system. The values currently used in theoretical modeling studies (Lu et al., 2008; Mitchell et al., 2005; Mitchell et al., 2008) were chosen empirically in order to reproduce the steady-state prevalence values observed in different herds. An observational study found an increased risk of infection from infected dams, higher shedding prevalence, or the presence of an infected calf, but the values provided by this were in terms of odds ratios rather than rates (Benedictus et al., 2008). It may be of interest to consider a Bayesian time-to-event model, which would estimate a constant or time-dependent infectious pressure based on the time between birth and positive test results, incorporating existing knowledge of compartment transition rates and test sensitivity.

Similar models have been developed for clinical outcomes of disease in swine (Baadsgaard et al., 2004).

In summary, the Bayesian method proposed in this study was insufficient for accurately estimating MAP transmission rates in endemically infected commercial dairy herds. Other estimation methods should be explored.

# REFERENCES


Auranen, K., E. Arjas, T. Leino, and A. K. Takala. 2000. Transmission of pneumococcal carriage in families: a latent Markov process model for binary longitudinal data. J. Am. Stat. Assoc. 95(452):1044-1053.

Baadsgaard, N. P., S. Hojsgaard, Y. T. Grohn, and Y. H. Schukken. 2004. Forecasting clinical disease in pigs: comparing a naive and a Bayesian approach. Prev. Vet. Med. 64:85-100.

Becker, N., and T. Britton. 1999. Statistical studies of infectious disease incidence. J. R. Stat. Soc. Ser. B 61(2):287-307.

Benedictus, A., R. M. Mitchell, M. Linde-Widmann, R. W. Sweeney, T. L. Fyock, Y. H. Schukken, and R. H. Whitlock. 2008. Calculation of transmission parameters of Mycobacterium avium subspecies paratuberculosis infections in a dairy herd going through a control program. Prev. Vet. Med. 83(3-4):215-227.

Cauchemez, S., and N. M. Ferguson. 2008. Likelihood-based estimation of continuous-time epidemic models from time-series data: application to measles transmission in London. J. R. Soc. Interface 5:885-897.

Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, Jr., and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229(12):1912-1919.

Forrester, M. L., A. N. Pettitt, and G. J. Gibson. 2007. Bayesian inference of hospital-acquired infectious diseases and control measures given imperfect surveillance data. Biostatistics 8(2):383-401.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457-511.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter. 1996. Markov Chain Monte Carlo in Practice. Chapman & Hall, London.

Glass, K., N. Becker, and M. Clements. 2007. Predicting case numbers during infectious disease outbreaks when some cases are undiagnosed. Statist. Med. 26:171-183.

Green, P. J. 2003. Trans-dimensional Markov chain Monte Carlo. Pages 179-198 in Highly Structured Stochastic Systems. P. J. Green, N. L. Hjort, and S. Richardson, ed. Oxford University Press, Oxford.

Lanzas, C., P. Ayscue, R. Ivanek, and Y. T. Grohn. 2010. Model or meal? Farm animal populations as models for infectious diseases of humans. Nat. Rev. Microbiol. 8(2):139-148.

Lekone, P. E., and B. F. Finkenstädt. 2006. Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. Biometrics 62:1170-1177.

Lu, Z., R. M. Mitchell, R. L. Smith, J. S. Van Kessel, P. P. Chapagain, Y. H. Schukken, and Y. T. Grohn. 2008. The importance of culling in Johne's Disease control. J. Theo. Bio. 254:135-146.

Mitchell, R. M., S. M. Stehman, R. H. Whitlock, A. Benedictus, and Y. H. Schukken. 2005. A deterministic mathematical model of Mycobacterium avium subsp. paratuberculosis (MAP) transmission on commercial US dairy farms. in 8th International Colloquium on Paratuberculosis. Vol. Copenhagen, Denmark. S. S. Nielsen, ed., Copenhagen,Denmark.

Mitchell, R. M., R. H. Whitlock, S. M. Stehman, A. Benedictus, P. P. Chapagain, Y. T. Grohn, and Y. H. Schukken. 2008. Simulation modeling to evaluate the persistence of Mycobacterium avium subsp. paratuberculosis (MAP) on commercial dairy farms in the United States. Prev. Vet. Med. 83(3-4):360-380.

Neal, P. J., and G. O. Roberts. 2004. Statistical inference and model selection for the 1861 Hagelloch measles epidemic. Biostatistics 5(2):249-261.

O'Neill, P. D. 2002. A tutorial introduction to Bayesian inference for stochastic epidemic models using Markov chain Monte Carlo methods. Math. Biosci. 180:103-114.

O'Neill, P. D., and N. Becker. 2001. Inference for an epidemic when susceptibility varies. Biostatistics 2(1):99-108.

O'Neill, P. D., and G. O. Roberts. 1999. Bayesian inference for partially observed stochastic epidemics. J. R. Stat. Soc. Ser. A Stat. Soc. 162(1):121-129.

Raftery, A. E., and S. M. Lewis. 1992. One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. Statistical Science 7:493-497.

Rankin, J. D. 1961. The experimental infection of cattle with Mycobacterium johnei. III. Calves maintained in an infectious environment. J. Comp. Pathol. 71:10-15.

Roberts, G. O., and R. L. Tweedie. 1996. Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms. Biometrika 83(1):95-110.

Smith, R. L., Y. T. Grohn, A. K. Pradhan, R. H. Whitlock, J. S. Van Kessel, J. M. Smith, D. R. Wolfgang, and Y. H. Schukken. 2009. A longitudinal study for the impact of Johne's Disease status on milk production in individual cows. J. Dairy Sci. 92(6):2653-2661.

Smith, R. L., R. L. Strawderman, Y. H. Schukken, S. J. Wells, A. K. Pradhan, L. A. Espejo, R. H. Whitlock, J. S. Van Kessel, J. M. Smith, D. R. Wolfgang, and Y. T. Grohn. 2010. The effect of Johne's disease status on reproduction and culling in dairy cattle. J. Dairy Sci. 93:3513-3524.

Streftaris, G., and G. J. Gibson. 2004a. Bayesian analysis of experimental epidemics of foot-and-mouth disease. Proc. R. Soc. B 271:1111-1117.

Streftaris, G., and G. J. Gibson. 2004b. Bayesian inference for stochastic epidemics in closed populations. Statistical Modeling 4:63-75.

van Roermund, H. J. W., D. Bakker, P. T. J. Willemsen, and M. C. M. de Jong. 2007. Horizontal transmission of Mycobacterium avium subsp. paratuberculosis in cattle in an experimental setting: Calves can transmit the infection to other calves. Vet. Microbiol. 122(3-4):270-279.

van Schaik, G., C. R. Rossiter, S. M. Stehman, S. J. Shin, and Y. H. Schukken. 2003. Longitudinal study to investigate variation in results of repeated ELISA and culture of fecal samples for Mycobacterium avium subsp paratuberculosis in commercial dairy herds. Am. J. Vet. Res. 64(4):479-484.

Whitlock, R. H., B. L. Mangold, S. McAdams, T. L. Fyock, R. W. Sweeney, Y. H. Schukken, J. M. Smith, J. S. Van Kessel, E. Hovingh, J. S. Karns, D. R. Wolfgang, and T. E. Johnson. 2007. Real-time PCR testing of pooled (1:5) fecal samples comparison to HEYM culture. in .

Whitlock, R. H., R. W. Sweeney, J. M. Smith, J. S. Van Kessel, E. Hovingh, J. S. Karns, D. R. Wolfgang, T. L. Fyock, S. McAdams, and Y. H. Schukken. 2008. Cattle shedding MAP: a new paradigm, passive shedding or active shedding? Pages 12 in Johne's Disease Integrated Program 4th Annual Conference. USDA-CSREES-NRI-CAP, East Lansing, MI.

Whitlock, R. H., S. J. Wells, R. W. Sweeney, and T. J. Van. 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): sensitivity and specificity of each method. Vet. Microbiol. 77(3-4):387-398.

APPENDIX 1

STOCHASTIC MODELING OF INDIVIDUAL ANIMALS FOR SPREAD OF *MYCOBACTERIUM AVIUM SUBSP. PARATUBERCULOSIS* IN ENDEMICALLY INFECTED DAIRY HERDS

***Introduction***

In order to simulate data for validation of the Markov Chain Monte Carlo (MCMC) model, 2 stochastic models were developed. Each produces individual animal information from endemically infected dairy herds, based on a user-defined set of parameters.

***Models and Methods***

These models are based on the age-stratified map model first proposed by Mitchell et al. (Mitchell et al., 2008), shown in figure A.1.1. Briefly, animals are classed into 3 age groups (calves, <1 year old: heifers, 1-2 years old; adults, >2 years old) and grouped according to infection status (X, susceptible; Tr, transiently shedding; H, latent; $Y_1$, low-shedding; $Y_2$, high-shedding).

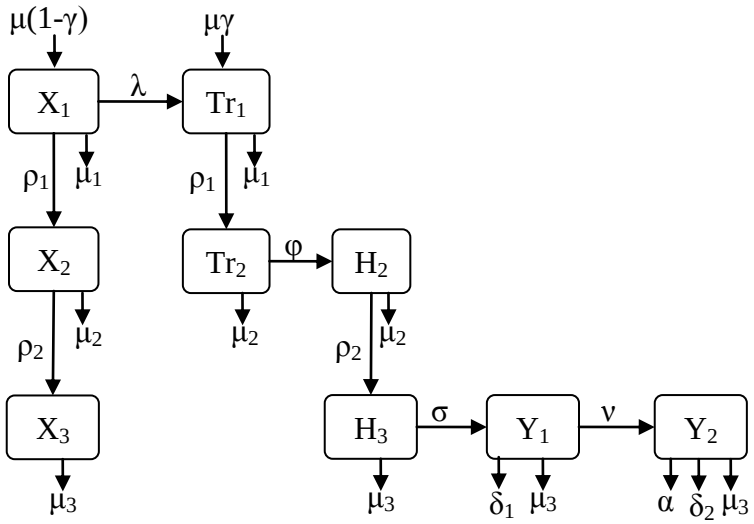



Figure A.1.1: A schematic diagram of the conceptual model used to simulate animal data. All classes and parameters are defined in the text.

Calves are born at rate μ to be susceptible ($X_1$), or vertically infected ($Tr_1$) with probability γ. Susceptible calves can be infected at rate λ to become transient shedders ($Tr_1$) or can age into resistant heifers ($X_2$) at rate $\rho_1$. Resistant heifers age into resistant adults ($X_3$) at rate $\rho_2$. Transiently shedding calves age at rate $\rho_1$ into transiently shedding heifers ($Tr_2$), which become latent heifers ($H_2$) at rate φ. Latent heifers become latent adults ($H_3$) at rate $\rho_2$, and latent adults become low-shedding adults ($Y_1$) at rate σ. Low-shedding adults become high-shedding adults ($Y_2$) at rate ν, and high-shedding adults are culled for clinical disease at rate α. Each age class experiences a general mortality rate (calves, $\mu_1$; heifers, $\mu_2$; adults, $\mu_3$). This model can be defined by the following system of ordinary differential equations (ODE):

**Equation A.1.1**

$$\frac{dX_1}{dt} = \mu(1-\gamma)N - (\mu_1 + \rho_1 + \lambda)X_1$$

$$\frac{dTr_1}{dt} = \mu\gamma N + \lambda X_1 - (\mu_1 + \rho_1)Tr_1$$

$$\frac{dX_2}{dt} = \rho_1 X_1 - (\mu_2 + \rho_2)X_2$$

$$\frac{dTr_2}{dt} = \rho_1 Tr_1 - (\mu_2 + \varphi)Tr_2$$

$$\frac{dH_2}{dt} = \varphi Tr_2 - (\mu_2 + \rho_2)H_2$$

$$\frac{dX_3}{dt} = \rho_2 X_2 - \mu_3 X_3$$

$$\frac{dH_3}{dt} = \rho_2 H_2 - (\mu_3 + \sigma)H_3$$

$$\frac{dY_1}{dt} = \sigma H_3 - (\mu_3 + \nu + \delta_1)Y_1$$

$$\frac{dY_2}{dt} = \nu Y_1 - (\mu_3 + \alpha + \delta_2)Y_2$$

where

**Equation A.1.2**

$$\gamma = \frac{\gamma_1(H_3 + Y_1) + \gamma_2 Y_2}{N}$$

$$\lambda = \beta_1(Tr_1 + Tr_2 + Y_1) + \beta_2 Y_2$$

$$\mu = \frac{\mu_1(X_1 + Tr_1) + \mu_2(X_2 + Tr_2 + H_2) + \mu_3(X_3 + H_3 + Y_1 + Y_2)}{N}$$

$$N = X_1 + Tr_1 + X_2 + Tr_2 + H_2 + X_3 + H_3 + Y_1 + Y_2$$

and all parameter definitions and values are given in Table A.1.1. As a reference, this model was simulated in the *lsoda* function of the *odesolve* package in R.

*Gillespie Algorithm*

The first model presented used Gillespie's direct algorithm to simulate stochastic dynamics of MAP transmission (Keeling and Rohani, 2008). In this method, the rate of each possible event is calculated; the sum of all rates is then the total rate of change of the system. The time to the next event is calculated as an exponentially distributed random variable with a rate equal to the total rate of change. The event to occur is drawn at random from all possible changes, weighted by their relative rates. For the MAP model, values for parameters are presented in Table A.1.1 and the possible events and their individual rates are presented in Table A.1.2. This model is based on the stochastic modeling system presented by Lu et al. (Lu et al., 2010). Herd size was set originally to 171, with 100 adult animals, 33 heifers (between 1 and 2 years of age), 33 susceptible calves (<1 year old), and 5 transiently shedding calves.

Table A.1.1: Parameter values for the compartmental model of MAP

| Parameter | Definition | Value (/year) | Reference |
|---|---|---|---|
| $\mu_1$ | death rate of calves | 0.111 | 1 |
| $\mu_2$ | death rate of heifers | 0.007 | 1 |
| $\mu_3$ | death rate of adults | ODE, tau-leap: 0.33<br>Gillespie: $\mu_{3inv}+\mu_{3v}$ | 1<br>7 |
| $\mu_{3inv}$ | involuntary death rate of adults | 0.048 | 1 |
| $\mu_{3vol}$ | voluntary death rate of adults without disease-based culling | 0.285 | 1 |
| $\mu_{3dz}$ | disease-based culling rate of adults | $\delta_1 Y_1/N_3+(\delta_2+\alpha)Y_2/N_3$ | 7 |
| $\mu_{3v}$ | voluntary death rate of adults with disease-based culling | $\mu_{3vol}-\mu_{3dz}$ | 7 |
| $\mu_b$ | female birth rate | 0.37 | 1 |
| $\delta_1$ | culling rate of low-shedders | 0 | 8 |
| $\delta_2$ | culling rate of high-shedders | 0 | 8 |
| $\alpha$ | culling rate of high-shedders due to clinical disease | 0.7 | 2 |
| $\rho_1$ | rate of aging from calf→heifer | 1 | 9 |
| $\rho_2$ | rate of aging from heifer→adult | 1 | 9 |
| $\varphi$ | rate of transition from transient shedding to latency | 2 | 3 |
| $\sigma$ | rate of transition from latency to low shedding | 0.667 | 2 |
| $\nu$ | rate of transition from low shedding to high shedding | 0.33 | 2 |
| $\gamma_1$ | probability of vertical transmission due to latency or low shedding | 0.15 | 4 |
| $\gamma_2$ | probability of vertical transmission due to high shedding | 0.17 | 4 |
| $\beta_1$ | infection rate of transient and low shedders | 0.002, 0.01, 0.05 | 8 |
| $\beta_2$ | infection rate of high shedders | $4\beta_1$ | 7 |
| $Se_1$ | sensitivity of fecal culture to diagnose low-shedding animals | 0.5 | 5 |
| $Se_2$ | sensitivity of fecal culture to diagnose high-shedding animals | 0.9 | 6 |

References: 1: (NAHMS, 1996); 2: (Whitlock et al., 2000) and (van Schaik et al., 2003); 3: (Rankin, 1961); 4: (Whitlock et al., 2005); 5: (Whitlock et al., 2000); 6: (Collins et al., 2006); 7: calculated; 8: user-defined; 9: assumed

Table A.1.2: Events and their rates for the stochastic model based on the Gillespie algorithm

| **Event** | **Change** | **Rate** |
|---|---|---|
| Death of $X_1$ | $X_1 \rightarrow X_1\text{-}1$ | $\mu_1 X_1$ |
| Death of $X_2$ | $X_2 \rightarrow X_2\text{-}1$ | $\mu_2 X_2$ |
| Death of $X_3$ | $X_3 \rightarrow X_3\text{-}1$ | $\mu_3 X_3$ |
| Death of $Tr_1$ | $Tr_1 \rightarrow Tr_1\text{-}1$ | $\mu_1 Tr_1$ |
| Death of $Tr_2$ | $Tr_2 \rightarrow Tr_2\text{-}1$ | $\mu_2 Tr_2$ |
| Death of $H_2$ | $H_2 \rightarrow H_2\text{-}1$ | $\mu_2 H_2$ |
| Death of $H_3$ | $H_3 \rightarrow H_3\text{-}1$ | $\mu_3 H_3$ |
| Death of $Y_1$ | $Y_1 \rightarrow Y_1\text{-}1$ | $\mu_3 Y_1$ |
| Death of $Y_2$ | $Y_2 \rightarrow Y_2\text{-}1$ | $\mu_3 Y_2$ |
| Cull $Y_1$ | $Y_1 \rightarrow Y_1\text{-}1$ | $\delta_1 Y_1$ |
| Cull $Y_2$ | $Y_2 \rightarrow Y_2\text{-}1$ | $\delta_2 Y_2$ |
| Clinical cull of $Y_2$ | $Y_2 \rightarrow Y_2\text{-}1$ | $\alpha Y_2$ |
| Progression of $Tr_2$ | $Tr_2 \rightarrow Tr_2\text{-}1$; $H_2 \rightarrow H_2\text{+}1$ | $\varphi Tr_1$ |
| Progression of $H_3$ | $H_3 \rightarrow H_3\text{-}1$; $Y_1 \rightarrow Y_1\text{+}1$ | $\sigma H_2$ |
| Progression of $Y_1$ | $Y_1 \rightarrow Y_1\text{-}1$; $Y_2 \rightarrow Y_2\text{+}1$ | $\nu Y_1$ |
| Susceptible birth from $X_3$ | $X_1 \rightarrow X_1\text{+}1$ | $\mu_b X_3$ |
| Susceptible birth from $H_3$ | $X_1 \rightarrow X_1\text{+}1$ | $(1\text{-}\gamma_1)\mu_b H_3$ |
| Susceptible birth from $Y_1$ | $X_1 \rightarrow X_1\text{+}1$ | $(1\text{-}\gamma_1)\mu_b Y_1$ |
| Susceptible birth from $Y_2$ | $X_1 \rightarrow X_1\text{+}1$ | $(1\text{-}\gamma_2)\mu_b Y_2$ |
| Infected birth from $H_3$ | $Tr_1 \rightarrow Tr_1\text{+}1$ | $\gamma_1 \mu_b H_3$ |
| Infected birth from $Y_1$ | $Tr_1 \rightarrow Tr_1\text{+}1$ | $\gamma_1 \mu_b Y_1$ |
| Infected birth from $Y_2$ | $Tr_1 \rightarrow Tr_1\text{+}1$ | $\gamma_2 \mu_b Y_2$ |
| Direct transmission by $Tr_1$ | $X_1 \rightarrow X_1\text{-}1$; $Tr_1 \rightarrow Tr_1\text{+}1$ | $\beta_1 Tr_1 X_1$ |
| Direct transmission by $Tr_2$ | $X_1 \rightarrow X_1\text{-}1$; $Tr_1 \rightarrow Tr_1\text{+}1$ | $\beta_1 Tr_2 X_1$ |
| Direct transmission by $Y_1$ | $X_1 \rightarrow X_1\text{-}1$; $Tr_1 \rightarrow Tr_1\text{+}1$ | $\beta_1 Y_1 X_1$ |
| Direct transmission by $Y_2$ | $X_1 \rightarrow X_1\text{-}1$; $Tr_1 \rightarrow Tr_1\text{+}1$ | $\beta_2 Y_2 X_1$ |

To produce individual animal data, the initial herd at time 0 was simulated by randomly sampling birth dates from a reasonable range of years for each group of animals, with years divided into tenths: calves from time 0 to time -1, heifers from time -1 to time -2, and adults from time -2 to time -7. Infection time for transiently infected calves was set at time -1. For each animal, the following information was recorded in a permanent file: birth date ($t_b$), death date ($t_d$), infection date ($t_I$), date of latency onset ($t_H$), date of shedding onset ($t_{Y1}$), date of high-shedding onset ($t_{Y2}$), date of first low-positive test ($t_{low}$), and date of first high-positive test ($t_{high}$). Any dates

corresponding to events that did not occur for an individual were recorded as 0. Current category status was also tracked for each animal, with possible statuses of $X_1$, $X_2$, $X_3$, $Tr_1$, $Tr_2$, $H_2$, $H_3$, $Y_1$, $Y_2$, or dead.

At the time of each event, an animal was selected from the eligible category for that event and the time of the event was recorded in that animal's history. For example, if a direct transmission occurred, an animal would be chosen at random from those animals classified as $X_1$ at time 1. The chosen animal would be assigned an infection time $t_I$=event time, and the animal's status would be changed to $Tr_1$. If the event was a birth, an animal was added to the herd with status $X_1$ and $t_b$=event time (if a susceptible birth) or status $Tr_1$ and $t_b=t_I$=event time (if an infected birth). If the event was a death, the oldest eligible animal was assigned a time of death $t_d$=event time and its status was changed to dead.

Two methods of animal aging were considered: stochastic and deterministic. For the stochastic aging model, the events in Table A.1.3 were added to the events listed in Table A.1.2 and aging was processed as any other event, with the oldest eligible animal moving to the next age category. For the deterministic aging model, it was assumed that animals do not age at random, so aging was handled in parallel to the random events modeled in the Gillespie algorithm. At each time step, after the random event was recorded, any animals that passed their first year of age during the time step (time-$t_b \geq 1$) were moved from $X_1$ to $X_2$ or from $Tr_1$ to $Tr_2$, as appropriate. Then, any animals that passed their second year of age during the time step (time-$t_b \geq 2$) were moved from $X_2$ to $X_3$, from $Tr_2$ to $H_3$, or from $H_2$ to $H_3$, as appropriate.

Table A.1.3: Additional events and their rates for the stochastic aging version of the model based on the Gillespie algorithm

| Event | Change | Rate |
|---|---|---|
| Aging of $X_1$ | $X_1 \rightarrow X_1\text{-}1$; $X_2 \rightarrow X_2+1$ | $\rho_1 X_1$ |
| Aging of $X_2$ | $X_2 \rightarrow X_2\text{-}1$; $X_3 \rightarrow X_3+1$ | $\rho_2 X_2$ |
| Aging of $Tr_1$ | $Tr_1 \rightarrow Tr_1\text{-}1$; $Tr_2 \rightarrow Tr_2+1$ | $\rho_1 Tr_1$ |
| Aging of $H_2$ | $H_2 \rightarrow H_2\text{-}1$; $H_3 \rightarrow H_3+1$ | $\rho_2 H_2$ |

After the full simulation, test results were calculated at annual intervals. Once in each year in the simulation, animals in category $Y_1$ at the testing date were sampled with probability $Se_1$; any sampled animals were assigned a low-positive test date $t_{low}$=time unless an earlier low-positive test result had been recorded. Likewise, on the same date, animals in category $Y_2$ were sampled with probability $Se_2$ and sampled animals were assigned a high-positive test date $t_{high}$=time unless an earlier high-positive test result had been recorded.

*Tau Leap*

The second model was simulated using the tau leap methodology (Keeling and Rohani, 2008). In this method, time is advanced in regular intervals; the number of times an event occurs over the time interval is calculated as a Poisson-distributed random variable with parameter rate*δt, where rate is the rate of the event and δt is the time interval. For the MAP model, infection pressure λ(t) at time t was calculated as

**Equation A.1.3**

$$\lambda(t) = \beta_1\big(Tr_1(t) + Tr_2(t) + Y_1(t)\big) + \beta_2 Y_2(t)\text{l}$$

where $\beta_1$ and $\beta_2$ are defined in Table A.1.1 and the compartments are defined above. The rate of infections, then, was $\lambda(t)*X_1(t)$. After the number of infection events was drawn from the Poisson distribution for a time point, the animals to be infected were

chosen at random from all susceptible animals at that time point, $X_1(t)$, and assigned an infection time, $t_I=t$.

All other events were calculated from exponential distributions based on the parameters described in Table A.1.1. At birth, a calf was assigned a date of death based on $\mu_1$; if that date was greater than the time at which the animal entered the heifer compartment, $t_{age1}$, the animal survived to become a heifer, the first death date was discarded and a new date of death was assigned based on $\mu_2$; if the second death date was greater than the time at which the animal entered the adult compartment, $t_{age2}$, the animal survived to adulthood and a third date of death was assigned based on $\mu_3$. In mathematical terms,

**Equation A.1.4**

$$t_d = \begin{cases} t_{d,calf} & if \quad t_{d,calf} < t_{age1}, t_{d,calf} \sim Exp(\mu_1) \\ t_{d,heifer} + t_{age1} & else, if \quad t_{d,heifer} < t_{age2}, t_{d,heifer} \sim Exp(\mu_2) \\ t_{d,adult} + t_{age2} & else \quad t_{d,adult} \sim Exp(\mu_3) \end{cases}$$

where $t_d$ is the time of death. The time of aging was modeled both stochastically and deterministically. In the stochastic aging model, for each animal, $t_{age1} \sim Exp(\rho_1)$ and $t_{age2} \sim Exp(\rho_2)+t_{age1}$, with the restriction that $t_{age1} \leq 24$ months and $t_{age2} \leq 36$ months. In the deterministic aging model, $t_{age1}=1$ year and $t_{age2}=2$ years for all animals. The number of calves born in a timestep was set equal to the number of animals dying in that timestep; calves were born at time t with vertically-transmitted infection with probability $\gamma(t)$, where

**Equation A.1.5**

$$\gamma(t) = \gamma_1\big(H_3(t) + Y_1(t)\big) + \gamma_2 Y_2(t)$$

and all parameters are defined in Table A.1.1.

Infected animals were assigned transition times (transient →latent, latent →low-shedding, low-shedding →high-shedding) based on the exponential distributions with the rates defined in Table A.1.1. All transitions with times after the assigned time of death were assumed not to occur. Animals that become high-shedding were assigned a new death date with rate $\mu_3+\alpha$. All transition times were assigned at the time of infection. For each animal, then, a permanent file is maintained with the following information: birth date ($t_b$), death date ($t_d$), date of entering heifer group ($t_{age1}$), date of entering adult group ($t_{age2}$), infection date ($t_I$), date of latency onset ($t_H$), date of shedding onset ($t_{Y1}$), date of high-shedding onset ($t_{Y2}$), date of first low-positive test ($t_{low}$), and date of first high-positive test ($t_{high}$).

The model proceeded with the following algorithm:

1. Time is updated, $t=t+\delta t$
2. Newborn calves are added, based on number of deaths during previous time step, and vertical infections are added based on $\gamma(t)$.
3. The number of animals in each category (Figure A.1.1) is updated, and $\lambda(t)$ is calculated.
4. The number of horizontal infections is calculated, and transition times are assigned for each newly infected animal.
5. The number of animals in each category (Figure A.1.1) is updated, and $\gamma(t)$ is calculated.

An initial herd was simulated with 100 newborn animals, and steps 1 and 2 were repeated for 120 months to establish stable population dynamics. At that point, the 5 youngest calves were assigned an infection time of 120 and all transition times were

assigned. Steps 1 to 5 were then repeated until the time reached 110 years; simulations in which fadeout occurred were discarded, as only endemic dynamics were of interest.

*Comparisons*

Test results were simulated post-hoc for the individual animal records; all animals were assumed to be tested on an annual basis with the fecal culture test. The number of animals testing positive on a testing date was simulated as a binomial distribution with the number of animals shedding at a given level and the probability of testing positive being the sensitivity of fecal culture for animals at that shedding level. The first date at which an animal tested positive at a level (low or high) was recorded.

All three models (ODE, Gillespie, and tau leap) were simulated for a set of 3 transmission rates for low-shedding animals ($\beta_1$), assuming that high-shedding animals were 4 times as infectious ($\beta_2=4\beta_1$). All were simulated for an initial herd with 166 animals. For the Gillespie and ODE models, this consisted of 100 adults, 33 heifers, and 33 calves, 5 of which were transiently infected at the start time. For the tau leap model, the herd was initialized with 166 calves, herd dynamics were stabilized at 120 months, and the 5 youngest calves were infected. The simulation was continued for 100 years after introduction of infection to ensure endemicity; fadeouts were rare (0-2 per 100 realizations) and were discarded.

All models were programmed and simulated in R 2.11.1 (R Development Core Team, 2010), which was accessed through the Revolution R Analytics interface (© 2010 Revolution Analytics, Inc.).

***Results***

All models reached endemic equilibrium within 25 years. The deterministic endemic prevalence of the ODE model and the distribution of prevalence at endemicity for 15 realizations each of the Gillespie and tau leap models are shown in Figure A.1.2. The distribution of test-positive animals present monthly in 15 realizations each of the Gillespie and tau leap models are shown in Figure A.1.3.

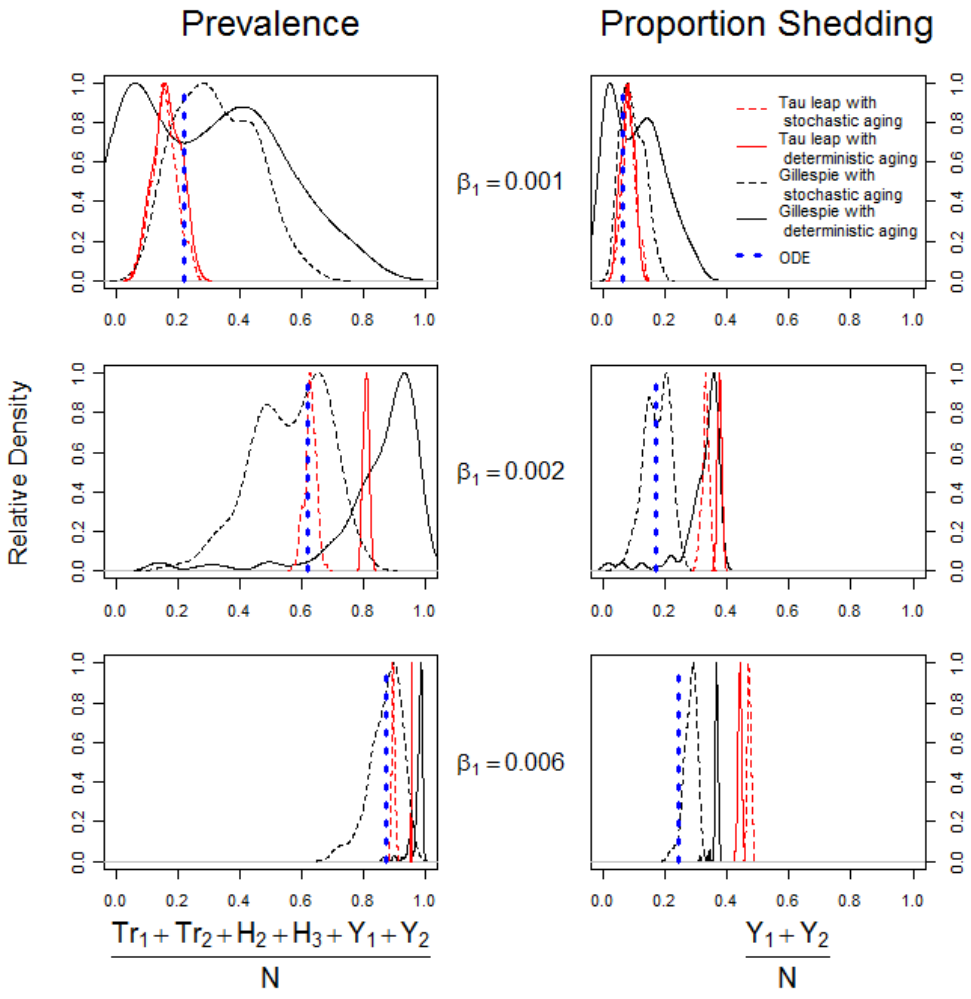



Figure A.1.2: Comparison of distribution of mean endemic prevalence for 100 realizations each of the Gillespie and tau leap stochastic models with stochastic and deterministic aging, compared to the endemic prevalence from the deterministic (ODE) model, over 3 levels of transmission rate. All herds had approximately 100 adult animals and were simulated for 100 years.

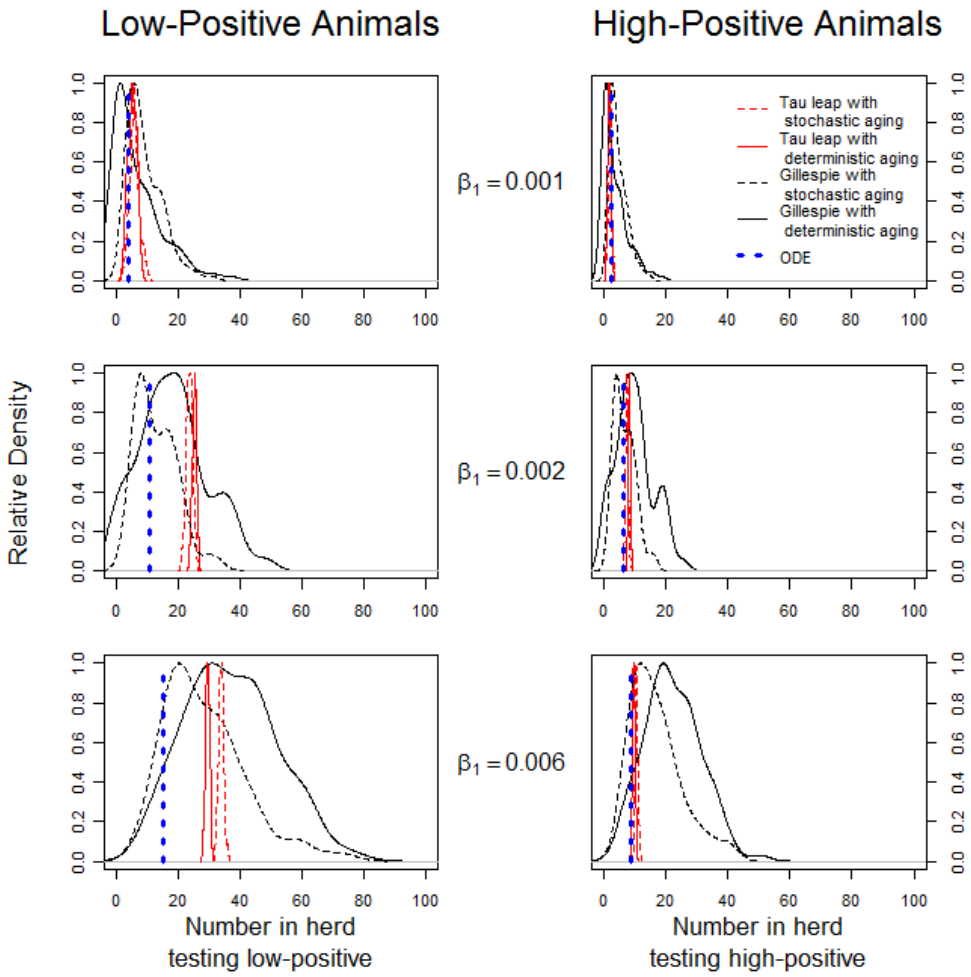



Figure A.1.3: Comparison of distribution of average number of test-positive animals present each month at the endemic state for 100 realizations each of the Gillespie and tau leap stochastic models with deterministic and stochastic aging, over 3 levels of transmission rate. All herds had approximately 100 adult animals and were simulated for 100 years.

The results of the Gillespie model with stochastic aging after 1000 simulations are shown in Figure A.1.4; a greater number of simulations decreases the amount of stochastic effect in the distributions, allowing for the true distribution to be seen more clearly.

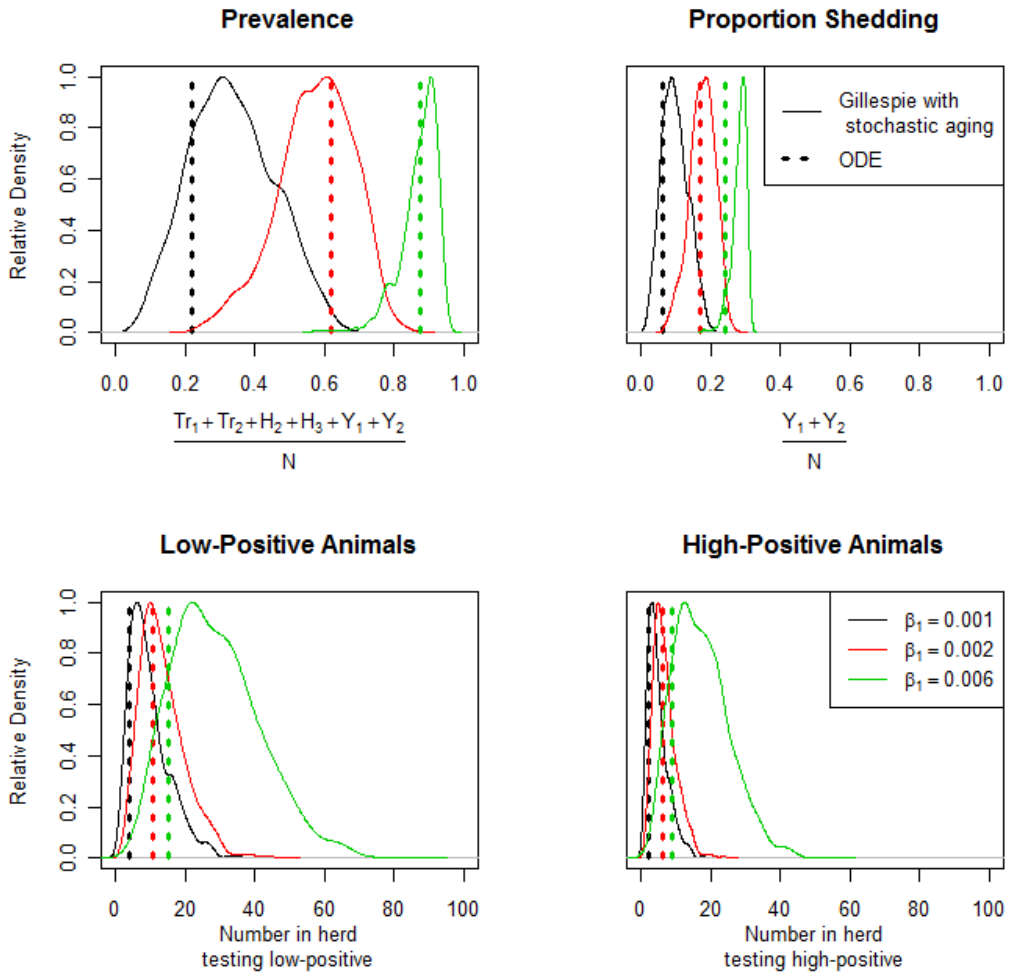



Figure A.1.4: Comparison of the mean endemic prevalence, proportion of shedding animals, number of low-positive animals, and number of high-positive animals over 1000 simulations of the Gillespie model with stochastic aging to the number predicted by the ODE model over 3 levels of transmission rate $\beta_1$. All herds had approximately 100 adult animals and were simulated for 100 years.

***Discussion***

Three models, one deterministic and two stochastic, were presented based on a similar understanding of MAP dynamics in dairy herds. The stochastic models had 2 different variations each for simulating aging. These different methodologies produced similar results, but with some variations.

The tau leap methodology is an approximation of Gillespie's direct algorithm, and the results shown here illustrate the effect of the approximation. The tau leap model was more sensitive to the transmission rate than the Gillespie model with stochastic aging, as was the Gillespie model with deterministic aging. As expected, most results of the ODE model most closely resembled the results of the Gillespie model with stochastic aging, which used the most similar assumptions about the system. The tau-leap method uses a timestep greater than the average timestep in the Gillespie direct algorithm. This allows multiple changes to occur during each timestep, and these changes are not independent. So long as multiple changes do not occur to the same animal during a single timestep, this lack of independence in not a fatal flaw; however, the impact may still be observed in the results. For all outputs considered, the tau leap simulations showed less variability than the Gillespie simulations, and all models produced more variability when transmission rates were low. Gillespie's direct algorithm has more inherent room for variability than the tau leap methodology, especially in this model, as it only allows a single event to happen at any timestep. At low transmission rates, the lumpiness of the system (changes must occur on the level of whole animals) is more likely to increase the stochastic effect for all models, especially for small herd sizes such as considered here. Figure A.1.4 shows that, despite the variability in results, the mean of the Gillespie model with stochastic aging truly follows the predicted results in the ODE model.

The results of the Gillespie model with deterministic aging also do not resemble the results of the ODE model so well as that with stochastic aging; the difference in assumptions about the duration of susceptibility allows for higher

prevalence at high transmission rates. As no animals gain resistance before 1 year of age, more infections become possible; with stochastic aging, animals that aged out of the calf strata later were more likely to be infected.

All of the stochastic models predict a higher number of low-positive animals than the ODE model at high transmission rates. The ODE model prediction is based on the probability of testing positive given a single test. The stochastic models, on the other hand, simulate annual testing of animals, and animals that have previously tested positive retain that status, regardless of future test results. Thus, an animal remaining in the herd for more than one year after a positive test result can increase the total number of test-positive animals in those future years, even if the animal were to test negative at all future times. This effect is stronger for low-shedders and at high transmission rates because higher numbers of low-shedding animals are tested in general and higher numbers of shedding animals are tested in high transmission rate herds (with the corresponding high prevalence). Testing more shedding animals increases the number of animals with previous test-positive results.

In all models, changes to the transmission rates produce a greater variation in the true prevalence than in the proportion of animals shedding or in the observed prevalence. While true prevalence is sensitive to the transmission rates, observed prevalence is rather insensitive. This is due to the limitations of the diagnostic method: sensitivity to low-shedding animals is poor and sensitivity to latent animals is 0. As latent and low-shedding animals comprise the bulk of the infected adults, and as only adults are tested for MAP, a large proportion of infected animals will remain

undetected. This leaves a very small range of test-positive animals, which does not allow for a great deal of variability to be observed.

The purpose of the stochastic models presented here is to produce simulated datasets representing realistic animals in a dairy herd. The information produced for each animal can include only what is observable in a herd (birth and death date, date of first low- and high-positive fecal culture) or can be expanded to include information we would like to observe (i.e., dates of infection and onset of shedding). These data will be of use in validating statistical models designed for field data.

## REFERENCES


Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, Jr., and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229(12):1912-1919.

Keeling, M. J., and P. Rohani. 2008. Modeling Infectious Diseases in Humans and Animals. Princeton University Press, Princeton, NJ.

Lu, Z., Y. H. Schukken, R. L. Smith, and Y. T. Grohn. 2010. Stochastic simulations of a multi-group compartmental model for Johne's disease on US dairy herds with test-based culling intervention. J. Theo. Bio. 264(4):1190-1201.

Mitchell, R. M., R. H. Whitlock, S. M. Stehman, A. Benedictus, P. P. Chapagain, Y. T. Grohn, and Y. H. Schukken. 2008. Simulation modeling to evaluate the persistence of Mycobacterium avium subsp. paratuberculosis (MAP) on commercial dairy farms in the United States. Prev. Vet. Med. 83(3-4):360-380.

NAHMS. 1996. NAHMS 1996: Johne's Disease on U.S. Dairy Operations. Pages 1-58 in . USDA:APHIS:VS, ed. National Animal Health Monitoring Survey, Fort Collins, CO.

R Development Core Team. 2010. R: A language and environment for statistical computing. in R Foundation for Statistical Computing, Vienna, Austria.

Rankin, J. D. 1961. The experimental infection of cattle with Mycobacterium johnei. III. Calves maintained in an infectious environment. J. Comp. Pathol. 71:10-15.

van Schaik, G., Y. H. Schukken, C. Crainiceanu, J. Muskens, and J. A. VanLeeuwen. 2003. Prevalence estimates for paratuberculosis adjusted for test variability using Bayesian analysis. Prev. Vet. Med. 60(4):281-295.

Whitlock, R. H., S. J. Wells, R. W. Sweeney, and T. J. Van. 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): sensitivity and specificity of each method. Vet. Microbiol. 77(3-4):387-398.

Whitlock, R. H., M. Widmann, R. W. Sweeney, T. L. Fyock, A. Benedictus, R. M. Mitchell, and Y. H. Schukken. 2005. Estimation of parameters on the vertical transmission of Map in a low-prevalence dairy herd. in 8th International Colloquium on Paratuberculosis. Vol. Copenhagen, Denmark. S. S. Nielsen, ed., Copenhagen, Denmark.

# CONCLUSION

## DISCUSSION AND FUTURE AVENUES OF RESEARCH

The research presented in this dissertation has increased knowledge of *Mycobacterium avium* subsp. *paratuberculosis* (MAP) and Johne's disease (JD) biology, effects, and control in US dairy herds. However, there are many paths open for further research, some opened by the results of this research.

Chapter 1 examined the effect that JD status has on milk production. The results show that high-shedding animals have decreased average milk production and that the milk production of animals that test positive decreases over time, as the disease progresses. The study also found higher average milk production in animals that will test MAP-positive, compared to animals that never test positive. One possible explanation for this is a genetic link between high milk production and susceptibility to MAP or JD; this could be further studies with association studies for MAP infection status and genes known to affect milk production.

Chapter 2 looked at the effect of JD status on time to culling and on calving intervals. The results of the culling analysis show that test-positive animals are culled at a faster rate than test-negative animals, which may be the result of the decreasing milk production observed in Chapter 1 or of the efforts to control MAP in the study herds. The results of the calving analysis show that high-shedding animals have lower calving rates than test-negative animals, but low-shedding animals have higher calving rates than both groups. Including culling probability in the calving analysis critically changed the results of the analysis in low-shedding animals, indicating that future

time-to-event studies of dairy reproduction should take into account the dependent nature of culling as a censoring method.

Chapter 3 attempted to analyze the presence of MAP in the dairy environment and to assess the association between MAP in the environment and test status of the cattle in the herd. The results show that culture of environmental samples is not a sensitive method for detecting MAP in a dairy herd; however, the amount of MAP in a pen-level composite manure sample, over time, can be indicative of the number of high-shedding animals housed in that pen. As the amount of MAP in the environment was related to the average amount of MAP shed by cattle in the herd, control strategies that remove high-shedding animals from the herd should be effective in decreasing the environmental load.

Chapter 4 presents a statistical model for estimating MAP transmission rates, using data that are readily available in many dairy herds. Unfortunately, the amount of information available is limited by the sensitivity of the diagnostic tests, with the result that the model is unable to accurately predict the transmission rate from simulated data. Other methods of determining the transmission rate will be necessary.

The initial purpose of the research presented here was to provide parameters necessary for economic modeling of MAP control in commercial dairy herds. The impact of JD status on milk production and reproduction directly impacts the net present value (NPV) of a dairy cow, but the change in culling rates is most likely in response to that change in NPV. In discrete-time economic models based on the compartmental model of MAP transmission, the distinction is not important and culling rates can be modeled based on JD status. In agent-based models, however, a

decreased NPV would automatically raise the culling rate in JD-positive animals. Thus, the application of these results will change depending on the format of the economic model. In contrast, the transmission rate results could be used in the same way by any economic analysis that considers secondary infections.

The application of this research in economic modeling has begun, and 2 manuscripts are currently under review. A compartmental model has been developed to optimize NPV under a given MAP control strategy, either by manipulating the herd-wide culling rate or by vaccination in combination with culling. This model can identify the MAP control strategy with the highest long-term NPV and/or the shortest time to fadeout. Unfortunately, this model can only deal with averages, assuming homogeneity within compartments, and is not equipped for stochasticity. Future research into stochastic optimization would be of interest.

Others have considered the economic cost of MAP on dairy farms, and have found that the cost of uncontrolled MAP could be quite large. The impact of milk production has been found to have the largest marginal effect on the cost of JD (Ott et al., 1999; Raizman et al., 2009), although decreased slaughter value and the loss of future value in infected cows have been included as well (van Schaik et al., 1996). A field study of 6 herds implementing individually-designed control programs found that the category responsible for most MAP-associated economic losses was milk production (Pillars et al., 2009). Over the course of a lactation, test-positive animals have been found to average less milk income and lower salvage value (Raizman et al., 2009), as we would predict based on the results presented here.

We have found in the research reported here that the impact of MAP on milk production and reproduction is limited for subclinical animals, suggesting that identifying these animals through testing may not be necessary to maintain high production levels. Instead, prevention of infection and removal of clinically infected (high-shedding) animals may be the best control strategy. Preliminary results of the economic compartment model utilizing these results show that hygiene management is the most cost-effective strategy for controlling MAP (J. Cho, personal communication). This has been found repeatedly by a number of different studies and models. Considering a variety of test-and-cull and hygiene strategies, the JohneSSim model found that hygiene management was a cost-effective strategy and that test-and-cull often was not (Groenendaal and Galligan, 2003), and the PTB-Simherd model also found that optimal hygiene management was better at controlling prevalence and increasing milk production and slaughter value (Kudahl et al., 2007). A similar model for MAP control in beef herds found that test-and-cull added little to hygiene management improvements (Bennett et al., 2010). Projecting an observed decline in MAP-associated losses over a 20-year time scale, Pillars et al. (2009) found that herds with established MAP control strategies would decrease their NPV by adding test-and-cull to their existing management plans. It is apparent from these models and from our results that removal of low-shedding (subclinical) animals from a well-managed dairy herd is not necessary.

One of the impacts disease control has on the economics of dairy production is the reaction of consumers to the disease (Chi et al., 2003). While consumers are currently unconcerned with the presence of MAP in dairy herds, the issue may become

more important in the future; there are some researchers positing a role for MAP in the etiology of Crohn's Disease (CD), a human disorder that superficially resembles JD (Mendoza et al., 2009; Uzoigwe et al., 2007). Animals shedding MAP in feces also frequently shed MAP in milk, and MAP has been cultured from commercial, pasteurized milk (Eltholth et al., 2009). In the case that a link between MAP and CD becomes more established, control of MAP in dairy herds will become a public health and public relations concern (Groenendaal and Zagmutt, 2008). In that eventuality, control of MAP may be mandated (with limits to the amount of MAP in bulk tank milk being the most likely regulation) or encouraged through bonus payments for milk and heifers from MAP-free herds. Under mandates, the economic model described above would need a further constraint, related to the amount of MAP shed in milk by low and high shedding animals. Under bonus payments, the model would need a change in net revenue (for milk) and sale value (for heifers) in herds meeting the testing requirements to be considered MAP-free (USDA:APHIS:VS, 2010). In either case, the optimization process would likely recommend different control programs.

It is important to note, however, that all results discussed here presume well-managed dairy herds. There is a selection bias present in most on-farm studies of disease, in that producers willing to enroll in these programs are often more proactive and aware of disease control measures. These herds often start studies with low prevalence and high hygiene levels. The impact of test-and-cull may be much higher in herds with high prevalence, especially in the early stages of MAP control. The agent-based model for MAP control should be able to predict the optimal length of a testing program in herds in which that program would be beneficial; no existing

models have considered testing strategies with a time span shorter than the length of the projection.

It is also important to note that the true impact of hygiene management programs is difficult to assess. These are aimed at decreasing transmission rates, or infectious pressure, between adults and calves. However, until the current study, direct estimation of MAP transmission rates has been limited. The model presented in Chapter 4 is unable to estimate transmission rates in herds, which could then be compared based on hygiene management strategies to establish a quantitative effect of such strategies on transmission, but other methods have been suggested and may prove useful. However, this requires a large number of herds with a high variety of strategies to make any statistically meaningful comparison; it also requires that hygiene strategies remain fixed for at least 5 years in order to estimate impact properly. In the future, we will be developing new methods to analyze the efficacy of hygiene strategies.

The results above can provide us with several general guidelines for cost-effective management of MAP in dairy herds. In general, good hygiene management should be more effective than test-and-cull programs at reducing the losses due to MAP in well-managed dairy herds. Removal of high-shedding animals is important, however, due to high infectiousness and low milk production and calving rates; these animals are frequently culled quickly on well-managed herds due to milk production losses. Regular environmental testing of adult cow pens, while ineffective for determining MAP infection status, could also be used to identify the presence of high-shedding animals to be removed.

In conclusion, while the work presented within this dissertation has elucidated several areas of MAP biology and control, there are many avenues yet to be explored. This research will provide a foundation for 2 branches of further research: economic optimization of MAP control and statistical modeling of mycobacterial diseases in humans and animals. Moving forward, we will continue to explore both of these avenues.

# REFERENCES

Bennett, R., I. McClement, and I. MacFarlane. 2010. An economic decision support tool for simulating paratuberculosis control strategies in a UK suckler beef herd. Prev. Vet. Med. 93:286-293.

Chi, J., A. Weersink, J. A. Van Leeuwen, and G. P. Keefe. 2003. The economics of controlling infectious diseases on dairy farms. Can. J. Ag. Econ. 50:237-256.

Eltholth, M. M., V. R. Marsh, S. Van Winden, and F. J. Guitian. 2009. Contamination of food products with Mycobacterium avium paratuberculosis: a systematic review. J. Appl. Microbiol. 107(4):1061-1071.

Groenendaal, H., and D. T. Galligan. 2003. Economic consequences of control programs for paratuberculosis in midsize dairy farms in the United States. J. Am. Vet. Med. Assoc. 223(12).

Groenendaal, H., and F. J. Zagmutt. 2008. Scenario analysis of changes in consumption of dairy products caused by a hypothetical causal link between Mycobacterium avium subspecies paratuberculosis and Crohn's Disease. J. Dairy Sci. 91:3245-3258.

Kudahl, A. B., S. Ostergaard, J. T. Sorensen, and S. S. Nielsen. 2007. A stochastic model simulating paratuberculosis in a dairy herd. Prev. Vet. Med. 78(2):97-117.

Mendoza, J. L., R. Lana, and M. Diaz-Rubio. 2009. Mycobacterium avium subspecies paratuberculosis and its relationship with Crohn's disease. World J. Gastroenterol. 15(4):417-422.

Ott, S. L., S. J. Wells, and B. A. Wagner. 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. Prev. Vet. Med. 40(3-4):179-192.

Pillars, R. B., D. L. Grooms, C. A. Wolf, and J. B. Kaneene. 2009. Economic evaluation of Johne's disease control programs implemented in Michigan dairy farms. Prev. Vet. Med. 90(3-4):223-232.

Raizman, E. A., J. P. Fetrow, and S. J. Wells. 2009. Loss of income from cows shedding Mycobacterium avium subspecies paratuberculosis prior to calving compared with cows not shedding the organism on two Minnesota dairy farms. J. Dairy Sci. 92:4929-4936.

USDA:APHIS:VS. 2010. Uniform Program Standards for the Voluntary Bovine Johne's Disease Control Program. in USDA-Animal and Plant Health Inspection Service-Veterinary Services, Washington,D.C.

Uzoigwe, J. C., M. L. Khaitsa, and P. S. Gibbs. 2007. Epidemiological evidence for Mycobacterium avium subspecies paratuberculosis as a cause of Crohn's disease. Epidemiol. Infect. 135:1057-1068.

van Schaik, G., C. H. J. Kalis, G. Benedictus, A. A. Dijkhuizen, and R. B. M. Huirne. 1996. Cost-benefit analysis of vaccination against paratuberculosis in dairy cattle. Vet. Rec. 139:624-627.