Modelling Coxiella burnetii spread within a dairy cattle herd

Aurélie Courcoul

N° d'ordre : 4258 **ANNÉE 2010**

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1
Label « Doctorat Européen »

Mention : Biologie

Ecole doctorale Vie Agro Santé

présentée par

Aurélie Courcoul Lochet

préparée à l'Unité Mixte de Recherche 1300 BioEpAR
Bio-agression, Epidémiologie et Analyse de Risque en Santé Animale
Oniris-INRA

Modélisation de la propagation de *Coxiella burnetii* en troupeau bovin laitier

Modelling *Coxiella burnetii* spread within a dairy cattle herd

Thèse soutenue à Nantes
le 10 Décembre 2010

devant le jury composé de :

Hans HEESTERBEEK
Professor, Utrecht University (the Netherlands) / *rapporteur*
Ann LINDBERG
Associate Professor, National Veterinary Institute (Sweden) / *rapporteur*
Jean-Sébastien PIERRE
Professeur, Université de Rennes 1 / *examinateur*
Nicolas ROSE
Ingénieur de recherche, Anses / *examinateur*
Suzanne TOUZEAU
Chargée de recherche, INRA / *examinateur*
Elisabeta VERGU
Chargée de recherche, INRA / *co-encadrante de thèse*
François BEAUDEAU
Professeur, Oniris / *directeur de thèse*

ACKNOWLEDGEMENTS

The research presented in this thesis was performed at the department of Animal Health of the French National Institute for Agricultural Research (INRA), at the unit Oniris-INRA "Bioaggression, Epidemiology and Risk Analysis in Animal Health" (Nantes, France). It was supported by the following funding bodies:

- ✓ the French National Institute for Agricultural Research,
- ✓ the unit Oniris-INRA 1300 "Bioaggression, Epidemiology and Risk Analysis in Animal Health",
- ✓ the French Ministry of Foreign and European Affairs through the Van Gogh Programme.

I am very grateful for the financial support I received from the above entities, without which this research project would not have been possible.

I would like to express my sincere thanks to the members of my evaluation committee, Dr Heesterbeek, Dr Lindberg, Dr Pierre, Dr Rose and Dr Touzeau and to my two fantastic supervisors, Dr Beaudeau and Dr Vergu, who offered invaluable assistance, support and guidance throughout this study.

I also would like to express my sincere gratitude to my co-authors: Dr Denis, Dr Hogerwerf, Dr Klinkenberg, Dr Monod, and Dr Nielen, who greatly contributed to the work carried out and gave me a warm welcome in their teams.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF TABLES

Chapter 2

Chapter 3

Chapter 4

CHAPTER 1

GENERAL INTRODUCTION

Picture : A. Senkowski

Zoonoses are infections that are naturally transmitted between vertebrate animals and humans. Their importance is mainly linked to the danger they represent for humans. According to the World Health Organization (WHO), zoonoses strike down 14 million people around the world every year. At least 61% of all human pathogens are zoonotic, and have represented 75% of all emerging pathogens during the past decade [170]. Zoonotic diseases have always represented a risk for humans (e.g. rabies or anthrax are long-past known), but recent events, such as Bovine Spongiform Encephalopathy (BSE) or swine Influenza outbreaks, showed that major zoonotic diseases can have a huge economic and social impact and dominate the media headlines for some time. Moreover, zoonoses also prevent the efficient production of food of animal origin and create obstacles to international trade in animals and animal products.

Recently, the European Commission was concerned about the increase within the EU in the number of cases of Q fever, a zoonosis due the bacterium *Coxiella burnetii*. Although Q fever has been present in cattle, sheep and goats holdings since a long time [107], human cases were sporadically reported until 2007 and the infection was seen as a rare occupational disease for farmers, veterinarians, and slaughterhouse workers [42]. However, in 2008, a total of 1,594 confirmed cases were reported in the EU, mainly in the Netherlands and Germany, corresponding to a 165.5% increase compared with the number of confirmed cases reported in 2007 [39]. In 2009, there were more than 2,300 human cases in the Netherlands, mainly in the form of atypical pneumonia. 19.7% of them were hospitalized [159].

Q fever is essentially an airborne disease and human infection occurs mainly after inhalation of aerosols generated from excreta from infected livestock (abortion and birth material, faeces, urine, milk) [8]. In addition, in ruminants, reproductive disorders are frequent signs of infection [103] and can impact production and economic efficiency of the farm. Thereby, Q fever is an issue in both public and animal health. The control of this infection in ruminants is therefore crucial to limit the infection spread in livestock as well as the zoonotic risk. The European Food Safety Authority (EFSA) recently highlighted the need to objectively assess the relevant epidemiological parameters (such as rates of within-herd transmission, between-herd spread and spillover from animal populations to humans) and the effectiveness of control options for *C. burnetii* infection in domestic ruminant populations [39]. Our work is in line with the EFSA opinion.

We focused on dairy cattle and our aim was to better understand the within-herd pathogen spread in order to better control the infection.

I- Pathogen characteristics and host response to infection

In 1935, Australian scientists worked to identify the cause of a febrile illness among abattoir workers in Brisbane while American scientists struggled to identify a novel microorganism they had isolated from ticks. Without knowing it, they were working on the same pathogen, *C. burnetii*, a small obligate intracellular bacterium classified in the γ-subdivision of *Proteobacteria*, in the order of *Legionellales* [107].

Since its discovery, Q fever has been reported worldwide with the exception of New Zealand [56, 62]. The bacterium can infect a broad spectrum of hosts including domestic animals (livestock and pets), wildlife and even non-mammalian species including reptiles, fishes, birds and ticks. Infection occurs mainly after inhalation of contaminated aerosols. *C. burnetii* is extremely infectious: under experimental conditions, the inhalation of a single *Coxiella* cell can produce infection and clinical diseases in humans [150]. Within hosts, the bacterium resides within the phagolysosome of monocytes and macrophages. The organism may come in the form of a large cell variant (LCV), a small-cell variant (SCV), and a small dense cell (SDC). The LCV of *C. burnetii* is intracellular and metabolically active whereas the SDC and SCV forms are able to survive extracellularly as infectious particles [110]. The bacterium can then survive very well in the environment: up to 42 months at 4-6°C in milk, 12 to 16 months in wool, 120 days in dust, 49 days in dried urine and 30 days in dried sputum [115].

Several typing methods have been used for the characterisation of *C. burnetii* strains. Restriction fragment length polymorphism (RFLP) and pulsed field gel electrophoresis (PFGE) were performed for the differentiation of 80 *C. burnetii* isolates derived from animals and humans in Europe, USA, Africa and Asia. This allowed the distinction of 20 different groups corresponding to the geographical origin of the isolate. However, no correlation between restriction group and virulence of isolates was detected [68]. More recently, two Polymerase Chain Reaction based (PCR-based) methods have been described to type *C. burnetii*, MLVA (multi-locus variable number of tandem repeats analysis) [14, 147] and multispacer sequence typing (MST) [55]. To date, these techniques are considered to be the most discriminating methods for *C. burnetii*, allowing the identification of up to 36 distinct genotypes. In the near future, these tools will probably be very useful for epidemiological investigation, particularly to clarify linkages regarding the source of infection [39].

Cell-mediated immunity probably plays the critical role in eliminating *C. burnetii* [8]. According to Read et al. [129], the presence of either CD4+ or CD8+ T cells was sufficient to control infection in mice, and B cells were not necessary for primary immunity. However, other studies

suggested that both humoral and cell-mediated immune responses were important for host defense against *C. burnetii* infection: treatment of *C. burnetii* with immune sera was reported to make the bacterium more susceptible to phagocytosis and destruction by normal polymorphonuclear leukocytes or macrophages [171]. Moreover, *C. burnetii* exists in two antigenic phases called phase I and phase II. This phase variation phenomenon is similar to the smooth to rough lipopolysaccharide transition of other Gram-negative bacteria [12]. This antigenic difference is important in diagnosis in humans.

II- Q fever in Humans

1. Routes of transmission to humans

Airborne transmission of *C. burnetii* through inhalation of aerosolised bacteria or contaminated droplets and dust is the principal mode of transmission to humans [8]. Indeed, infected animals shed bacteria into the environment through faeces, vaginal mucus, urine, milk and especially parturition products [11, 20, 59]. Goats, sheep and cattle are recognized as the main source of human infection [96, 109, 142, 164] although dogs and cats can sometimes be involved in the transmission to humans [29, 124]. As *C. burnetii* survives very well in the environment, the bacterium contaminates aerosols and surrounding dust [167]. Moreover, wind plays a role in *C. burnetii* transmission: in France, a statistically higher incidence of human cases was associated with an increased frequency of the mistral one month before the onset of the disease [151]. In the literature, estimates regarding the distance that infectious particles can spread by air span a large range: from 400m up to 40 km, depending on the studies [38]. A recent study showed that, during the Dutch epidemics, persons living within 2 kilometres of an affected dairy goat farm had a much higher risk for Q-fever than those living more than 5 kilometres away [142]. Some environmental factors, soil moisture or vegetation density can also play a role in *C. burnetii* transmission [67].

Consumption of raw milk could be a source of infection [107]. However, according to the EFSA Panel on Biological Hazards, drinking milk containing *C. burnetii* can result in seroconversion but it remains unclear whether, and if so, to what extent, clinical disease can result from the consumption of milk or diary products, or of other foods containing *C. burnetii* [39]. Besides, the French Agency for Food Safety estimated that in case of human contamination by consumption of raw milk, the gross risk was to 'nil to negligible' [2].

Other transmission routes appear rare. Ticks can be naturally infected with *C. burnetii*, but they do not appear to be important in the maintenance of infections in humans [107]. Person-to-person or sexual transmissions are anecdotal [8, 112].

2. Clinical manifestations

Q fever is characterised by a clinical polymorphism and a frequently asymptomatic expression. Men are more often symptomatic than women, despite comparable exposure and seroprevalence, as well as people over 15 years compared to children [100, 152]. After an incubation period of approximately 20 days, the infection leads in around 40% of cases to an acute Q fever (Figure 1.1). The acute disease frequently includes fever, headaches, myalgias, arthralgias and cough [110]. Other manifestations are pneumonia, hepatitis, myocarditis, skin rash and neurologic signs [8, 128]. In acute cases of Q fever, the antibody level to phase II antigens is usually higher than the one of phase I, often by several orders of magnitude. In chronic disease, the reverse situation is observed.

Chronic Q fever may develop, many months to years after infection, in at-risk patients, i.e. patients with heart valve or vascular diseases or patients with cancer or immunosuppression. This chronic form appears in some 2% of acute symptomatic cases and the fatality rate may vary from 5 to 50% [38]. The most frequent manifestations are endocarditis and vascular infections but fever, loss of consciousness, weight loss, general fatigue, night sweats and hepatomegaly may also be present [168].

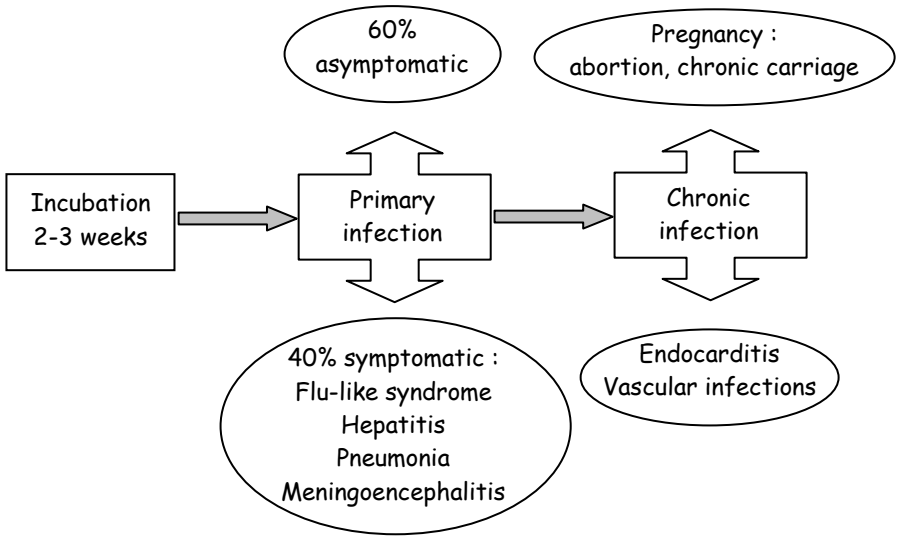

Figure 1.1. Q fever natural history in humans in the absence of treatment
(from Angelakis & Raoult [8])

In addition, when a pregnant women is infected by *C. burnetii* during pregnancy, there is a risk of abortion, preterm delivery or low birth weight [8].

3. Magnitude and distribution of human Q fever in the European Union

Q fever in humans is a communicable disease for which surveillance is mandatory in the EU. In 2007, 637 confirmed cases were reported to the European Centre for Disease Prevention and Control (ECDC), mainly in the Netherlands (168 cases), Spain (159 cases), Slovenia (93 cases), Germany (83 cases) and Bulgaria (36 cases). In 2008 and 2009, respectively 1,011 and 2,357 confirmed cases were reported in the Netherlands [39]. The months with the highest number of reported cases were July and August.

According to this spreading pattern, the EFSA concluded that human Q fever can be considered as a relatively infrequent clinical disease and that there is no obvious increase in the general disease risk [39]. However, human cases are likely to be underreported and the Q fever epidemic in the Netherlands has shown some divergence from the epidemic characteristics described until now: the infection persists over consecutive years and has sickened mainly people who never had contact with animals. The drivers of such an outbreak remain unclear. The epidemic could be caused by a more virulent strain of *C. burnetii* [42] or by changes of farm characteristics. According to the EFSA, most of the human Dutch Q fever cases are indeed linked to abortion in large dairy goats farms, and to a much lesser extent in dairy sheep farms [39].

III- Q fever in domestic ruminants

1. Modes of contamination

Inhalation of contaminated aerosols is the main route of infection for ruminants. Transmission by ticks is also possible: the bacterium was isolated in several tick species [63, 111, 154]. However, the importance of this mode of contamination has not been determined. In the same way, the contamination by ingestion of an infected placenta has, to our knowledge, not been evaluated yet. Although mice were found 10,000 times less susceptible to the infection when orally infected than when intraperitoneally infected [37], cats and dogs may be infected by the consumption of placentas [107]. Consequently, further studies are needed to quantify the

importance of this mode of contamination for ruminants. As viable *C. burnetii* were detected in the semen of bulls, sexual transmission of the bacterium between cattle should be possible [81], but the role of this route of transmission within a herd has not been explored until now. Lastly, *C. burnetii* infection has been reported in wild mammals (especially rats [166]) and birds which could represent a reservoir for *C. burnetii* infections.

2. Clinical manifestations

Most of the *C. burnetii* infected animals remain asymptomatic but reproduction disorders can occur [103]. In Plommet et al. [125], amongst eleven heifers inoculated by the intradermal route then inseminated, five aborted or remained sterile. Indeed, as the female uterus and mammary glands are primary sites of chronic infection [107], *C. burnetii* infection can induce abortions, stillbirth and delivery of weak lambs, calves and kids. In the majority of cases, abortion occurs at the end of gestation without specific clinical signs appearing before [12]. High abortion rates can be observed in some goat flocks [120, 135]. In cattle, metritis is frequently the unique manifestation of the disease [19, 153]. Aborting females recover rapidly and generally do not abort during the following gestations, while metritis can persist for several months [12]. Pneumonia and endocarditis are not described in animals except in experimental conditions. In Plommet et al. [125], all inoculated heifers developed a pneumonia 24 to 36 hours after the inoculation and 50% of the animals presented cardiac symptoms or pulmonary lesions in the months following the infection [125].

3. Characteristics of the bacterial shedding

C. burnetii infection in ruminants often becomes chronic, with persistent bacterium shedding: cows can shed *C. burnetii* for several months [59] and goats at successive parturitions [23]. This shedding is of major importance as it contaminates the environment and can lead to the infection of both susceptible animals and humans. For cows, ewes and goats, Rodolakis et al. [131] reported that, contrary to expectation, the shedding of *C. burnetii* could be not related to parturition. In *C. burnetii* infections, a great heterogeneity between shedders has been described [11, 37, 57, 131]: the shedding duration, level (i.e. the quantities of bacteria shed) and routes are variable between animals. Infected animals can indeed shed bacteria through birth products, vaginal mucus, faeces, and milk [132]. Amongst the three latter, no predominant route was identified in 242 dairy cows from 31 herds in which abortions due to *C. burnetii* were reported [57]. Besides, in the same study, 65% of the shedder cows shed by only one route. However, in asymptomatic herds, cows shed more frequently in milk than in vaginal mucus or faeces [59, 131]. The shedding duration and shedding levels are also variable

between animals: cows can shed from a sporadic to a three-month persistent way and the concentrations of bacteria shed in vaginal mucus or milk can vary from less than 100 bacteria/g to more than 1,000,000 b/g [59]. Although the presence of heavy shedder cows (i.e. animals shedding bacteria in higher titres and in a persistent way) was reported in milk [59], the role of these animals in *C. burnetii* transmission between animals and from animals to humans has not been determined yet. In goats and ewes, the same shedding routes are described and in the same way, are rarely concomitant [135]. Ewes were found to shed mostly in faeces and vaginal mucus [131] while goats were reported to shed mainly in milk and vaginal mucus [11, 131, 135].

The presence of heterogeneity in a population (e.g. variability in age, contact structure, infectiousness, etc…) is known to affect infection dynamics in many diseases. As an example, a model assuming that all farms and all animals are governed by the same underlying dynamics was unable to explain the highly overdispersed distribution of prevalences of *Escherichia coli O157* shedding on Scottish farms [106]. The best fit to the prevalence data was obtained when incorporating individual variability in transmission. In many cases, the heterogeneity of shedding has indeed a great impact on both the infection dynamics and the effectiveness of control measures: in dairy cattle infected by *Salmonella*, the presence of host heterogeneity in infectious period and contagiousness decreased the effectiveness of population-wide control strategies, making necessary the application of strategies targeting the most contagious animals [86]. For *C. burnetii* infections, the influence of this heterogeneity of shedding on the infection spread has not been evaluated yet. Hence, it is necessary to take into account the variability of the shedding duration, shedding levels and shedding routes in our epidemic model when representing *C. burnetii* spread and testing effectiveness of control strategies.

4. Diagnosis

Currently, the PCR is a sensitive and rapid mean to directly detect *C. burnetii* and therefore to identify the shedders [39]. This technique can be used on a wide range of samples (vaginal mucus, abortion material, faeces and milk). Real-time PCR is preferable to conventional PCR as it allows high sample throughput [122] and the quantification of the bacterium in the sample. As an example, a real-time PCR assay applied to bulk tank milk samples appears to be a valuable tool to assess on a larger scale the status of herds towards *C. burnetii* shedding [58]. Quantitative PCR kits are now commercially available. Moreover, although immunohistochemistry was until now useful when considering potential causes of abortions in

ruminants, multiplex PCR that can detect and differentiate between different abortive pathogens is now under development [24].

For the serological testing, the complement fixation (CFT) was considered the reference test for historical reasons [39]. However, the indirect immunofluorescence assay (IFA) and above all, ELISA, are now widely used. According to Kittelberger et al. [79], two commercial ELISA were more sensitive than the CFT in all panels from infected ruminants: their sensitivities were 81% for the Pourquier ELISA and 95% for the IDEXX ELISA. However, none of the tests are able to distinguish between acute and chronic infection or between vaccinated and naturally infected animals. Besides, a serological test does not give clear information about the individual infection status [39]: some infected animals are indeed seronegative while they are shedding.

5. Magnitude and distribution in the European Union

There are currently no EU rules about the notification and surveillance of *C. burnetii* infection and/or Q fever in domestic ruminants. Based on available data [39], *C. burnetii* is present in most, if not all, member states. It does not appear to be an increase in Q fever prevalence/incidence but comparison over time and between countries is problematic as there are considerable differences in testing protocol and data availability. In Gran Canaria island (Spain), 34.7% of 1,249 randomly selected ruminants (60.4% of goats, 31.7% of sheep and 12.2% of cattle) were reported seropositive using an indirect ELISA kit [134]. In northern Spain, a serosurvey was carried out in 1,379 sheep (42 flocks), 626 beef cattle (46 herds) and 115 goats (11 herds) [138]: ELISA anti-*C. burnetii* antibody prevalence was slightly higher in sheep (11.8 ± 2.0%) than in goats (8.7 ± 5.9%) and beef cattle (6.7 ± 2.0%); herd prevalence was 74% for ovine flocks, 45% for goat flocks and 43% for cattle. In Denmark, a study based on bulk tank milk samples from 100 randomly selected dairy herds demonstrated a prevalence of 59% antibody positive herds [3].

6. Control

In France, in infected herds, interventions against Q fever mainly consist in environmental measures such as destruction of placentas or disinfection of births locations, and in medical measures such as antibiotic treatment like injections of tetracyclines during the last month of gestation and vaccination [132].

As placentas and aborted foetuses contain high numbers of *C. burnetii*, births should preferentially take place in a specific location which can easily be disinfected and risk material

(placentas and aborted foetuses) should be collected and removed to specific rendering plants [2]. Moreover, the manure can be a potential vector of the infection [21], although exact levels of *C. burnetii* in manure have not been accurately determined yet. Either chemical [9] or heat treatment can be used to reduce the load of viable bacteria [39]. Tick control may also play a role in the transmission of the infection.

Observations concerning antibiotics are contradictory. In Berri et al. [22], oxytetracycline was administered in a flock of sheep after a Q fever episode but did not prevent further abortions and did not immediately suppress the shedding of the bacteria. However, this treatment may have affected the ewes in the long term, and prevented further spread of the infection to ewes and lambs. In Astobiza et al. [16], the oxytetracycline treatment neither prevented the shedding of bacteria nor limited the duration of bacterial excretion. The EFSA concluded that, although antibiotic treatment is used effectively in humans to reduce clinical symptoms associated with Q fever, the same treatment in animals is not effective in reducing neither the level nor the duration of *C. burnetii* shedding and should be avoided [39].

According to Rodolakis et al. [132], vaccination is an efficient tool to control the disease. Vaccination with an antigenic phase I vaccine in cattle was shown to suppress the shedding in milk, placenta and colostrum [25, 139]. More recently, Arricau-Bouvery et al. [13] compared the efficiency of phase I and phase II vaccines in goats: the phase I vaccine prevented abortions and dramatically reduced the frequency of bacteria shedding in the milk, vaginal mucus and faeces while the phase II vaccine did not affect the course of infection. Thus, phase I vaccines are much more effective than phase II vaccines. In Rousset et al. [136], the vaccine appeared neither able to prevent infection in exposed kids, nor to clear infection in infected goats, but was effective in reducing the level of shedding in a heavily infected herd. In fact, preventive vaccination (before infection) is much more effective than outbreak vaccination [39]. According to Guatteo et al. [61], a susceptible non pregnant cow had a five times lower probability of becoming a shedder than an animal receiving placebo. Vaccination seems a long-term control strategy but field and experimental data are needed to improve our understanding of the infection spread in and between infected vaccinated populations [39].

Other control options can be used in emergency situations when public health is at risk. Culling of pregnant animals, temporary breeding ban or control of animal movements are some of the measures implemented in the Netherlands during the current outbreak [39].

IV- Why a modelling approach to understand *C. burnetii* infection and assess control strategies?

1. Why not field observations?

To assess the effectiveness of control strategies within a herd, different approaches are available. Impacts of the control measures can be directly observed in the field. However, the spread of *C. burnetii* within a herd results in a complex process. The infection of susceptible animals is linked to the contamination of the environment, and therefore to the shedding of bacteria. There are different types of shedders (e.g. seronegative versus seropositive ones), and a same type of shedders shed by different shedding routes, for a variable duration and in variable quantities. Follow up the shedding within a herd is unfeasible on the long run. Moreover, if the aim is to compare different types of control strategies, such a follow-up would have to be performed in a high number of herds for a long period of time. A solution to monitor over time the spread of *C. burnetii* within a herd would be the use of a diagnostic test at the herd level. However, at the present time, although a real-time PCR applied to bulk tank milk samples is a valuable tool to assess the status of herds towards *C. burnetii* infection [60], there is no clear correlation between a positive result and either the prevalence of shedders in the herd or the environmental bacterial load. The direct monitoring of the environment is therefore of great interest. Recently in the USA, Kersh et al. [78] collected 1,600 environmental samples (mainly soil and dust on solid surfaces) and performed quantitative PCR. 23.8% of the samples analyzed were positive for *C. burnetii* DNA, and the locations that contained *C. burnetii* DNA were diverse: unsurprisingly dairy farms, cattle feed lots, veterinary hospitals, and goat-breeding facilities but also high schools, retail stores, grocery stores, football stadiums, banks, and post offices. Most of the samples analyzed had a fairly low number of bacteria detected but 10% of samples were much more contaminated. However, it is difficult to directly assess the viability of the bacteria in these samples: if the soil or dust sample is directly placed onto cultured host cells, the culture will be contaminated by a variety of microbes present in the sample. Viability is best determined by injection of environmental materials into mice but this test cannot be easily used in routine because of its cost and its logistical constraints. Thus, current methods monitoring the environmental bacterial load allow neither easily determining the risk of infection for susceptible individuals nor following the infection spread. A modelling approach seems then relevant when studying *C. burnetii* spread: the more complex a phenomenon is, or the more expensive and difficult it is to study, the more value there is to explore models [163].

2. Roles of epidemiological modelling

A model is a simplified representation of a complex phenomenon. By definition, all the models are "wrong" because they make simplifying assumptions [77]. However, modelling is an essential tool, particularly useful every time a major public health issue is raised. As an example, the infection by the Human Immunodeficiency Virus started being perceived in a different way when biomathematicians showed that observed data were compatible with the assumption that 100% of infected people develop the disease. Before, as only a small part of seropositive people showed clinical signs, the asymptomatic seropositive people were considered as healthy carriers [158].

In epidemiology, models have different roles. Prediction is the most obvious one and often aims at guiding policy decisions [77]. For quantitative prediction purposes, the model has to be accurate and validated (i.e. with the smallest uncertainty which could impact the conclusions). This objective is most of the time difficult to reach. When a new infection is introduced in a former susceptible area, no historical data are available. In addition, an epidemic reference situation in the absence of control measure or with a perfectly known control programme is rarely recorded, especially for animal infectious diseases [45]. Thus, most of the time, although models used to evaluate control strategies are sophisticated and parameter-rich, model conclusions are not quantitative. However, qualitative outputs are enough for a large range of purposes and especially for the comparison of different scenarios (e.g. spread of an infection in different regions, for different pathogen strains, for different control measures, etc…). A model also helps understanding how an infectious disease spreads in the real world [77]. It provides the modeller a virtual world in which everything can be recorded and every factor can be examined. For example, it is possible to explore the effects of variable numbers of partners on the spread of sexually transmitted diseases or the effects of neighbourhood contacts or animal purchases on the spread of livestock diseases. Besides, a model can allow estimating non observable parameters. For instance, some events can be very rare in the real world but have major consequences in public health. Their frequency has then to be assessed but this is hardly feasible through field observations. As an example, the residual risk of HIV infection through blood transfusion is now very low. It is unfeasible to assess by a comparative experiment the potential benefit of an additional prevention strategy. Instead, a model can be used to estimate this infection risk and simulate control scenarios [158]. Although estimation of parameters can be considered as a role in itself, it also helps understanding the infection process and is a prerequisite for prediction purposes. At last, modelling allows highlighting gaps of knowledge: to develop models, modellers need quantitative data whereas most of the

time, only qualitative data is available in the literature or known by experts. Models are then a mean to critically evaluate the range of information available on the modelled system.

It has to be highlighted that modelling and field or experimental work are complementary approaches (Figure 1.2). On one hand, as we will see later, data is required to conceptualize the model, estimate the value of parameters and validate the model. On the other hand, models help testing biological assumptions, optimizing experiments protocols or identifying gaps of knowledge.

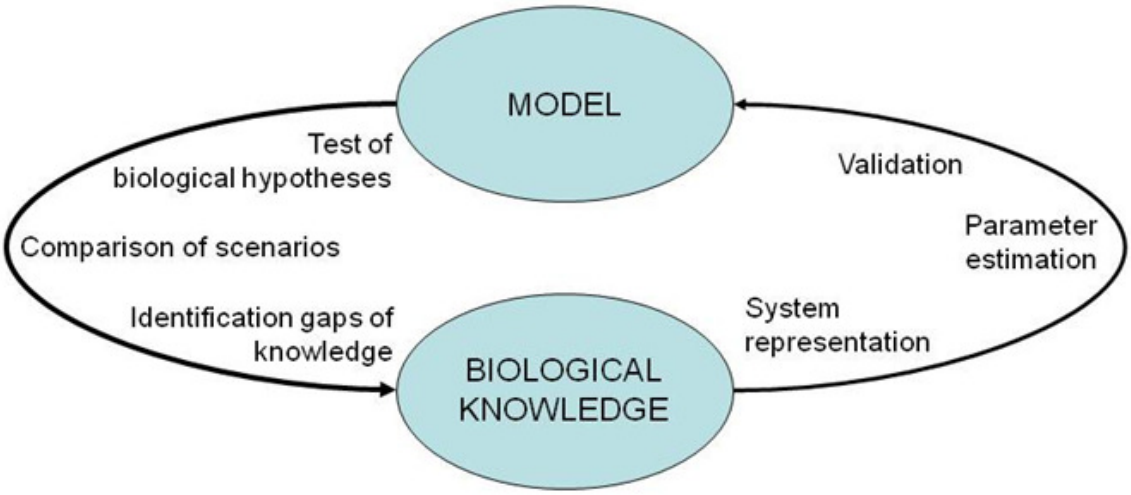

Figure 1.2. Mutual input of biology and modelling (from Ezanno et al. [45])

3. An example of epidemiological model in animal health

In livestock, epidemic models have been developed for various infectious diseases such as tuberculosis [71], brucellosis [40], BSE [6], Bovine Viral Diarrhoea [162], *Escherichia coli* infections [90, 157, 172], *Salmonella* infections [85], paratuberculosis [101], Contagious Bovine Pleuropneumonia [102], bluetongue [148] or foot-and-mouth disease [76].

A well-known example of mathematical model in domestic ruminants is the approach developed by Keeling et al. [74] representing the 2001 foot-and-mouth disease (FMD) outbreak. The aim of this study was to understand how spatial and individual heterogeneities influenced the course of the epidemic and to compare different vaccination and culling scenarios. The FMD outbreak was characterised by both a high probability of local spread and less frequent

longer-range transmission due to fomites and personnel movements. There was additional heterogeneity in farm size and species composition. In the model, the susceptibility and infectiousness of a farm were assumed varying according to its size and species composition. The model was fit to data, and a good agreement was obtained. It is worthy that the inclusion of the host species and herd-size heterogeneities in transmission were required to reach this good agreement and to properly simulate the spatial aggregation of cases. Large farms, and especially cattle ones, were indeed found to have a key role in the infection spread. Modeling demonstrated that culling infected farms, their direct contacts and contiguous farms was much more effective in reducing both the number of cases and the total number of culled farms than culling infected farms only. The delay from infection report to culling was also an important factor influencing the effectiveness of the control measures. Moreover, Keeling et al. considered the potential impact of both reactive and prophylactic vaccination on future FMD epidemics in the United Kingdom [75]. Mass prophylactic vaccination campaign could reduce the size and duration of the epidemic and vaccinating above 80,000 farms (over the 100,000 cattle farms in the UK) would even prevent almost all major epidemics. In addition, at the start of an outbreak, mass reactive vaccination, in combination with culling and animal movements restrictions might also control ongoing epidemics. On the contrary, ring vaccination would have a limited effectiveness.

This example shows (i) how important is to take into account the presence of host heterogeneity in a population, and (ii) how models can help understanding the course of the infection and guide decisions makers. From a practical point of view, this model played an important role in the formulation of the DEFRA's (Department of Environment, Food and Rural Affairs of the United Kingdom) contingency plan published in 2004 [77]. However, before model conclusions can inform policy, economical, sociological and logistical constraints have to be taken into account [75]. For example, farmers will have to be sure that vaccination will not devalue or limit the sale of their stock, or that they will get compensation. Besides, as vaccination will undoubtedly suppress clinical disease but not always infection, careful surveillance will therefore be required. Thus, model results have to be set back in the real world before being used for decision purposes. At last, this model was based on extensive data of a single epidemic. For an outbreak with a different strain (i.e. with different transmission properties and host specificities) or for an outbreak in another location (i.e. with different farming practices or weather conditions), the model will have to be adapted to the new situation before conclusions could be drawn from its outputs.

V- Objective and outline of the thesis

Because of the zoonotic risk induced by *C. burnetii* infections and their economic impact, there is a need to design effective control measures against *C. burnetii* spread in cattle. The general objective of this thesis is to develop a model representing the spread of *C. burnetii* within a dairy cattle herd in order to better understand the natural course of infection and to enlighten decision makers on the effectiveness of control measures. To provide a comprehensive description of the infection dynamics, we first aimed at estimating the main epidemiological parameters and then at identifying those that have the strongest impact on the disease spread pattern. When focusing on control measures, we aimed at comparing the effectiveness of different vaccination strategies in infected herds. Moreover, the identification of influencing parameters performed in the second part of our work could help to propose other potentially effective control strategies specifically impacting these key parameters of the disease spread.

Our general objective was reached in three stages (Figure 1.3). We first designed a model representing *C. burnetii* spread within a dairy herd and assessed its main epidemiological parameters from field data in a Bayesian framework. Secondly, as a great heterogeneity between *C. burnetii* shedders with a potential impact on the infection dynamics has been described, we chose to explicitly represent in our model the shedding routes and levels. We then performed a sensitivity analysis in order to identify the parameters, and especially those related to the heterogeneity of shedding, whose variation highly influences the infection dynamics. Lastly, we represented in the model different vaccination strategies and tested their comparative effectiveness by simulation.

Chapter 2 of the thesis first explains the main steps to set up a model. It then describes the field data we used to conceptualize the *C. burnetii* model and infer its parameters. The data set consists of individual health states of 235 cows of five chronically infected dairy herds sampled from one to five times over a four-week period. The stochastic individual-based model in discrete time we developed to represent the evolution of *C. burnetii* infection, as well as the Markov chain Monte Carlo methodology we used to estimate the parameters of interest are then presented.

Chapter 3 first details the importance of heterogeneity in a host population when studying an infection dynamics. It then describes the way the individual variability of the shedding duration, routes and levels were represented in the model. General aims and methods in sensitivity analysis are afterwards presented. Lastly, the approach based on a Principal

Component Analysis followed by an ANOVA that we performed is explained and the influence of the different epidemiological parameters on the model outputs is detailed.

Chapter 4 presents different ways to include vaccination in an epidemic model. It then describes how we adapted our dynamical model previously developed to simulate the impact of three different vaccination strategies (vaccination of both cows and heifers for 10 years, vaccination of heifers only for 10 years or vaccination of both cows and heifers for 3 years) on the infection dynamics and presents our results.

Finally, chapter 5 provides a general discussion on the PhD project. It presents the main results related to this thesis objectives and their potential field application. The modelling approach chosen is also discussed. Lastly, a few potential future directions are presented.

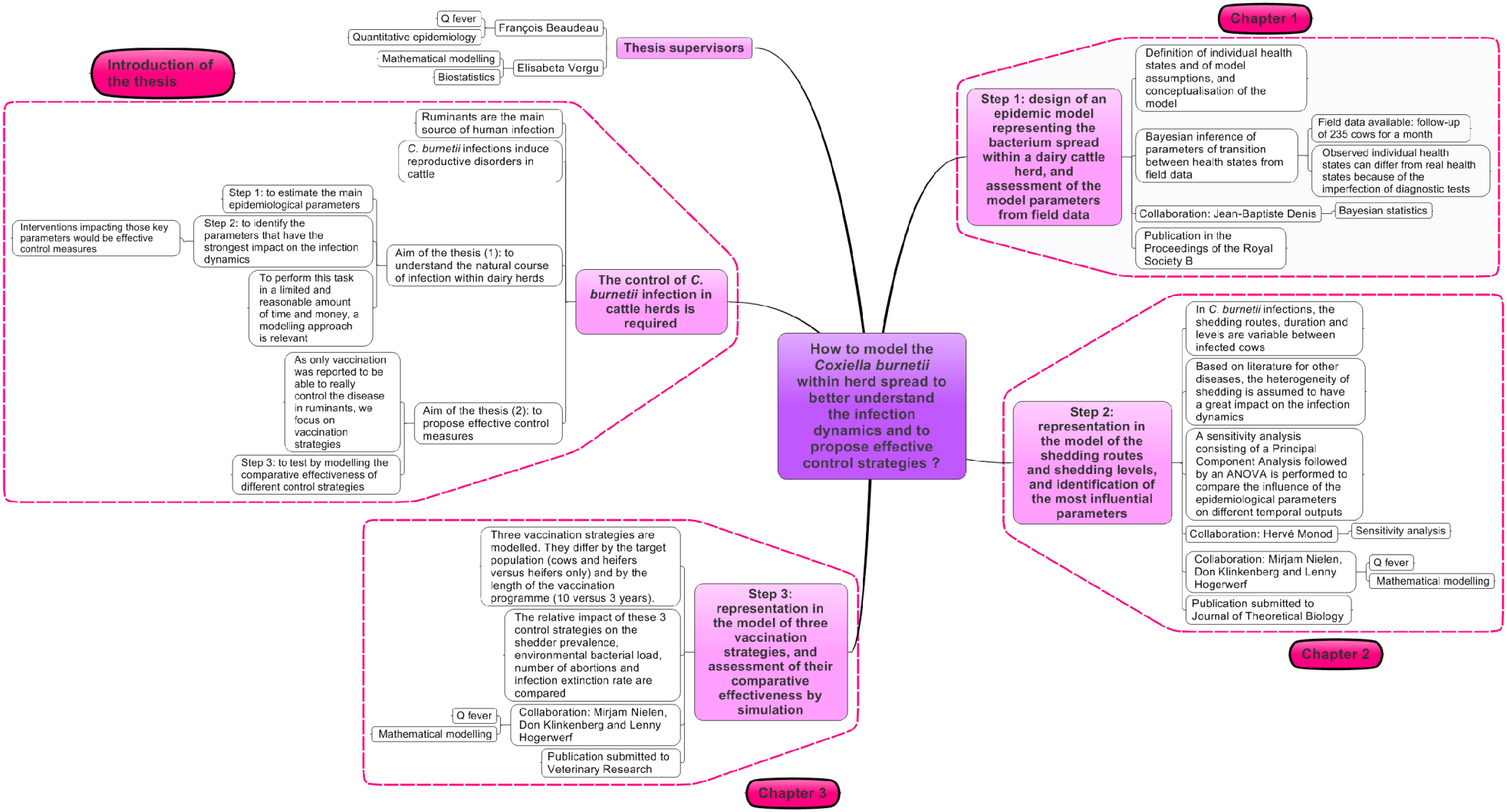

Figure 1.3. Mind map of the thesis project

CHAPTER 2

ELABORATION OF A MODEL REPRESENTING THE SPREAD OF *C. BURNETII* WITHIN A DAIRY HERD AND ESTIMATION OF ITS MAIN PARAMETERS FROM DATA

In the first section of this chapter, the main steps of the construction of a model and its confrontation to data will be described. Secondly, we will present the model we developed to represent the within herd *C. burnetii* spread. As both model conceptualization and parameter assessment depend on field data, we will first describe the data we used. Then, we will justify the model structure and formalism that we chose. At last, the main part of this chapter will be presented as it was published in the Proceedings of the Royal Society B [35]. It deals with the assessment of the main epidemiologic parameters involved in the dynamics of within herd *C. burnetii* spread from field data using a Bayesian approach.

I- Some generalities on how to build a model and to confront it to data

1. Choice of the model structure

The first step when developing a model for the spread of an infectious disease is to choose its structure: the different health states and transitions between them have to be defined. This backbone should reflect the natural history of the infection. If the population is not homogeneous with respect to the disease, the main categories in the population itself have also to be specified: according to Diekmann & Heesterbeek [36], the state of an individual is the set of information about the individual that is relevant to determine its future behaviour. It comprises its health state as well as other characteristics (such as age, genetic composition, stage of development, etc…) that may impact the infection dynamics. Classically, the health states considered are *S*, susceptible, and *I*, infectious. This *SI* model can be used for diseases like HIV: an individual is infectious as soon as infected, and for its whole life. An *SIS* structure is used for curable diseases: infected individuals are infectious until they are treated or recover and become susceptible again [163]. If there is a non negligible delay between the infection and the infectiousness, the health state *E*, exposed, can be added. Moreover, if individuals are immune to further infection after they have been infected, they enter the health state *R*, recovered. *SIR* and *SEIR* are standard model structures (Figure 2.1).

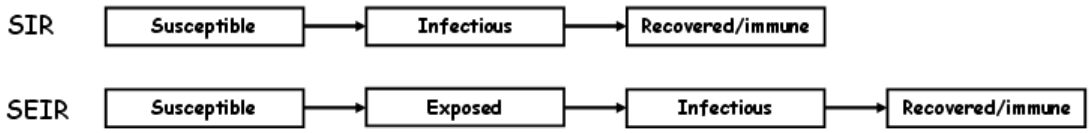

Figure 2.1. Structure of *SIR* and *SEIR* models

It has to highlight that the different phases of clinical and infection dynamics do not occur simultaneously (Figure 2.2). However, they are sometimes not distinguished in mathematical models (e.g., symptomatic and infectious periods are considered identical).

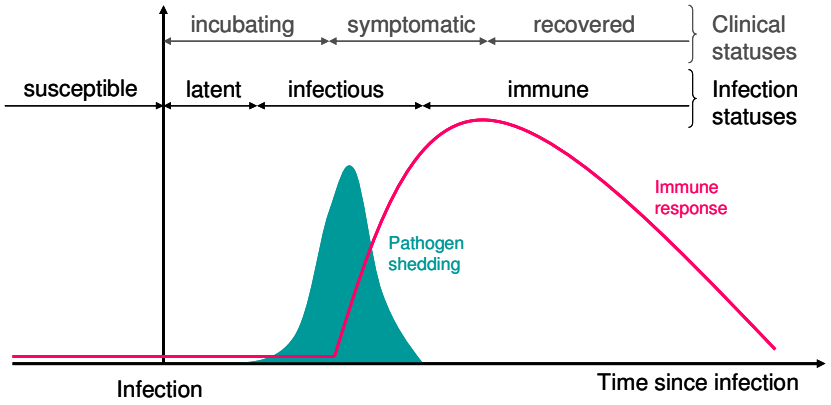

Figure 2.2. Individual infection status vs. clinical status in a simplified infection process (adapted from Keeling and Rohani [77] and Ezanno et al. [45]).

Depending on the disease, other health states and transitions can be specified: as an example, in some paratuberculosis models [101], health states T, transiently infectious, I_{S_low}, subclinically infected low shedder, I_{S_high}, subclinically infected high shedder or I_C, clinically affected, are represented. The definition of health states and their associated transitions is mostly based on the biology of the pathogen, host immune response and available data. The classic *SIR* structure can then be modified to obtain more sophisticated variants. In addition, the research question strongly determines the structure of the model [163]: as an example, if the model aims at exploring the impact of treatment use, considering additional states like 'successfully treated' or 'unsuccessfully treated' can be of great interest.

Besides, as previously mentioned, the population can be divided into distinct homogeneous classes with different behavioural characteristics (i.e. all members of a class have comparable risk of both contracting and transmitting infection) [77]. Age-structured models are frequently used when modeling childhood diseases, whereas risk-structured models are considered when modeling sexually transmitted diseases. For livestock diseases, models can

take into account the herd structure and different groups of animals (i.e. calves, heifers, lactating cows, and dry cows [40, 162]). The representation of host heterogeneity in models will be presented in more details in Chapter 3.

2. Choice of the mathematical formalism

Models can be deterministic, describing average dynamics, or stochastic, considering that chance can have a great impact on the infection dynamics [163]. As an example, in a *SIR* deterministic model, the rate at which individuals recover is fixed: for a given state of variable *I*, there is always the same number of individuals going from *I* to *R* during a time step. For a *SIR* stochastic model, *I* individuals have a given probability of transition from *I* to *R* and, due to random draws, the number of individuals going from *I* to *R* during a time step is variable from one model repetition to the other. Deterministic formulations are suited for large populations where randomness has relatively little overall impact, whereas stochastic models are more appropriate for small populations and rare events where the fluctuations have larger effects [45].

The scale and unit of modelling have also to be defined. The spread of a livestock infectious disease can indeed be modelled within a herd (the unit of modelling is then the animal) or between herds (the unit of modelling is then the farm). The scale is closely related to the research question. In most circumstances, disease transmission is a localized process. If, for instance, the study is aimed at exploring the infection spread in a school after an infected child is introduced, the unit and scale should be the individual and the population respectively. However, movements of individuals between human or animal populations facilitate the geographical spread of infectious diseases [77]. If the study is focused, for example, on determining the influence of neighbouring relationships and animal movements on the infection dynamics, a between herd scale should be considered in the model. For such models, the within farm infection dynamics is explicitly represented or not. In the previously described FMD model (see introduction IV.3.), the unit of modelling is the farm: due to the rapid transmission of the virus between animals situated at the same location, the within herd infection dynamics can be assumed negligible. Therefore, the whole farm is considered infected as soon as an animal is infected. On the contrary, for similar time scales but moderately spreading pathogens like the Bovine Viral Diarrhoea Virus, the within herd dynamics should be represented in details since it is unrealistic to assume that the entire farm is infected as soon as a single animal is infected [34]. For models describing the spread

of a pathogen at a larger scale (region, country, or world), the spatial positions of hosts are often taken into account. The previously cited FMD model aimed at simulating the spread of the FMD over the whole United Kingdom, it was a spatial model [74]: the locations of all farms were explicitly represented and the rate of transmission between two farms was expressed as a function of the distance.

Models can be compartmental or individual-based. In individual-based models, the health state of each individual is monitored over time, whereas compartmental models track the infection process for the individuals of a same health state collectively [163]. Individual-based models are often more computationally intensive but can also provide finer information. When host characteristics (such as age, sex, gestation status, intensity of contacts with the other individuals, etc.) are supposed to have an impact on the infection dynamics and are variable between individuals, it is more appropriate to develop individual-based approaches than compartmental models with many distinct compartments.

Concerning their time dependence, models can be in continuous time (the system could then be described by differential equations) or in discrete time (difference equations could then capture the dynamics). Differential equations provide a means for avoiding the issues regarding the size of the time step by describing events occurring continuously, rather than at discrete time intervals [163]. Indeed, in discrete time models, the choice of the time step is crucial: if this latter is too large (i.e. two successive transitions between health states can occur during a single time step), the model provides inaccurate and even non sense results. The appropriate size of the time step depends on the modelled phenomena: it should generally be less than the shortest average duration that individuals spend in a given health state [163].

At last, to describe the infection dynamics on the long run, key aspects of demography of the population considered (births, deaths, and migrations) may need to be incorporated in the model. In the case of animal populations being managed by humans (e.g. pigs or cattle herds), the representation of demography can be quite complex and sometimes require the development of an elaborated population dynamics model [44, 94].

3. Confrontation of the model to data

A second step after the model elaboration consists in determining appropriate and plausible values for model parameters. This can be done qualitatively based on information from the literature or expert's opinions. If data are available, model parameters can be quantitatively

assessed by fitting model predictions to data. However, sometimes this assessment is not feasible from existing knowledge or data and new data sets need to be collected and analysed using statistical methods. Although generic models can provide an intuitive explanation of the transmission of infectious diseases, it is only through detailed parameterization and rigorous assignment of numerical values to parameters that useful public health guidance can be generated [77].

When performing parameter inference from data different statistical techniques could be used. A widely known approach is that of "least squares". The sum of squares of the difference between the model predictions and the observed data is calculated in order to determine the parameter values which lead to the smallest value for this sum [163]. Another well-defined and widely-applied approach when fitting a model to data is that of "maximum likelihood": for a given set of parameters, the dynamics predicted by the model is determined. Then, the likelihood (i.e. the probability) that the observed data come from such dynamics is calculated. The best-fitting parameters are those which maximize this likelihood: the model is in closest agreement with the available data [77].

Bayesian statistical inference is also a widely-applied approach to assess parameter values. Although as for frequentist methods, the likelihood is still the key principle, there is an important difference in the way it is used. For a frequentist, parameter estimation is based solely on the likelihood while, for a Bayesian, it is based on both the likelihood and the prior information [99]. The prior distribution of a parameter is the probability distribution describing our initial knowledge about the parameter value. This distribution is based on previous studies or expert knowledge. This concept is criticized by frequentists as it introduces an element of subjectivity. In fact, estimated parameters values are an intermediate between observations and prior distribution and problems can occur when prior information is misleading and when one has a strong confidence in it [99]. The main difference is that frequentist statisticians consider model parameters as fixed but unknown while Bayesians consider them as random variables [130]. Bayesian methods often lead to more realistic estimated parameter values: a posterior distribution is obtained for each parameter, distribution which represents the uncertainty about the parameter, conditionally to data. Based on Bayes's theorem, the posterior distribution is expressed as:

$$\underbrace{p(\theta \mid D)}_{\text{Posterior distribution}} = \frac{\overbrace{p(D \mid \theta)}^{\text{Likelihood}}\ \overbrace{p(\theta)}^{\text{Prior distribution}}}{\int_{\theta} p(D \mid \theta)\, p(\theta)\, d\theta}$$ with D the Data and θ the parameters.

At this stage, it is impossible to determine the posterior distribution because the integral in the denominator is most often very difficult to obtain. However, this integral can be considered as a constant because it does not depend on the parameters θ. Markov chain Monte Carlo (MCMC) methods are then used: they are a class of algorithms which allow obtaining random draws from a probability distribution which is known up to a constant [123]. In our case, the prior distribution and the likelihood are known, so the numerator can be expressed. The posterior distribution is then known excepted for the denominator constant. Monte Carlo integration allows drawing samples from the target distribution and then calculating sample averages to approximate expectation. MCMC methods allow drawing these samples by appropriately constructing a Markov chain[1] that has the desired distribution as its equilibrium distribution [53]. The state of the chain after a large number of steps is then used as a sample from the desired distribution (i.e. the posterior distribution in our case). There are many algorithms designed for constructing these chains, but all of them, including the Gibbs sampler are special cases of the Metropolis-Hastings algorithm [53]. Several issues arise when implementing MCMC. A problem is to determine how many steps are needed to converge to the stationary distribution within an acceptable error. A good chain will have rapid mixing (i.e. the stationary distribution is reached quickly starting from an arbitrary position). The number of chains to be run, the starting values (to be chosen more carefully for slowly mixing chains), the length of burn-in (i.e. the first part of the chain to be removed in order to 'forget' the starting position) are also important technical adjustments to be considered in practice.

When confronting an SIR-like epidemiological model to the data, likelihood-based estimation of its parameters would be relatively easy to implement if the times of infection and removal were observed for all cases [32]. In practice, the transmission process is rarely completely observed (e.g. times of infection or removal are not observed for all individuals) and reported quantities may be aggregated (e.g. weekly). In this context, when the calculation of

[1] A Markov chain is a random process with the property that the next state depends only on the current state.

the likelihood becomes intractable, data augmentation methods (i.e. which allow augmenting the observed data with the missing information, for instance the times of infection or removal) were extensively used. MCMC sampling is of particular interest since it allows exploring the joint posterior distribution of parameters and augmented data. Although limited by the size of the augmented data (due to the computation times which dramatically increase with this size), MCMC approach is appropriate and efficient for small datasets (not exceeding a few thousands).

II- The modelling of *C. burnetii* spread within a dairy cattle herd

1. Description of the data set used for parameter estimation (data set A)

The data were collected by Raphaël Guatteo during its PhD and described in details in one of his papers [59]. R. Guatteo carried out a one-month longitudinal study in five French dairy cattle herds infected with *C. burnetii*, but without any clinical sign attributable to Q fever. The selected herds were chosen to satisfy two major criteria: (i) the presence of the bacterium *C. burnetii* within the herd; this was certified by a positive PCR result on bulk tank milk and more than 20% of cows seropositive for *C. burnetii*, and (ii) the absence of any control measure (i.e. antibiotics or vaccination directed against *C. burnetii*) before the end of the study. To assess the dynamics of *C. burnetii* infection, the lactating cows of these herds were sampled from one to five times on a weekly basis (Figure 1.3). The cows entering one of the herds during the study (as a consequence of a purchase or a first calving) were also included.

The individual state of each sampled cow was determined at each sampling time using an ELISA test (LSI ELISA Cox Ruminants®, Lissieu, France) on serum and a real-time PCR (LSI Taqvet *Coxiella burnetii*®, Lissieu, France) on three different samples (milk, faeces and vaginal mucus). The results of the ELISA test were expressed by the ratio (S/P) between optical densities of the sample and the positive control, and a cow was considered seropositive when the S/P ratio in serum was greater than or equal to 0.4. For the PCR test, only the samples presenting a typical amplification curve (demonstrating *C. burnetii* DNA

detection) with a Ct (cycle threshold) below 40 were considered positive. A cow was identified as PCR-positive when at least one of its three samples was PCR-positive.

At the initial point of the follow-up (day D0), the sizes of the five herds ranged from 24 to 79 lactating cows and a total of 217 cows were tested (Table 2.1). Thereafter, 100% of the initially (at D0) PCR-positive cows, 100% (or 50% in herds with more than 40 lactating cows) of the initially seropositive/PCR-negative cows, and 65% of the initially seronegative/PCR-negative cows were retained for the follow-up. Thus, during the following month, between 55% and 79% of the cows of each herd were tested every week (at D7, D14, D21 and D28) in the same way in order to determine their individual health state. According to the PCR results and ELISA test, at D0, between 35% and 74% of cows per herd were identified as PCR-negative/seronegative, between 1% and 23% were PCR-positive/seronegative, between 2% and 35% were PCR-positive/seropositive and between 17% and 37% were PCR-negative/seropositive. At the end point of the follow-up (D28), the herds comprised between 24 and 81 lactating cows.

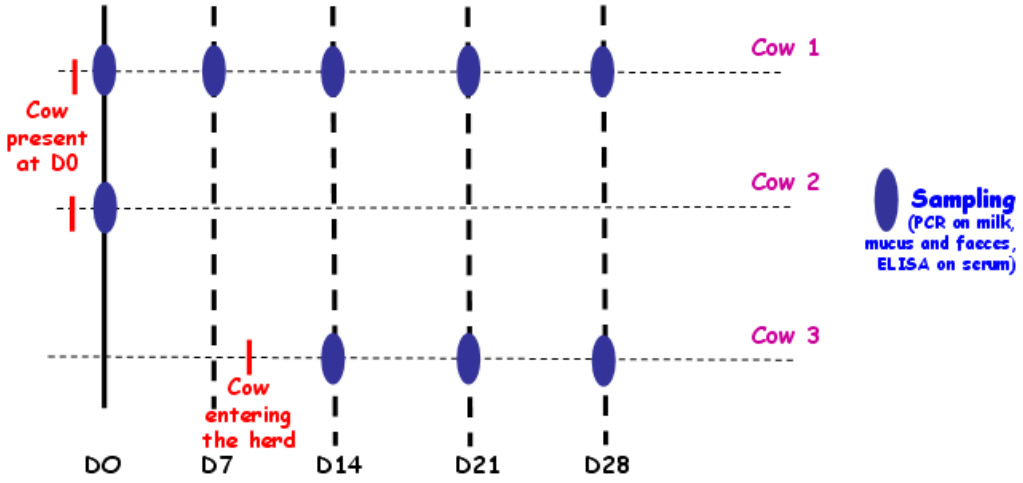

Figure 2.3. Sampling protocol during the one-month longitudinal study.
Cows present in the herd at time 0 of the follow-up were sampled either 5 times on a weekly basis or only once at time 0 (D0). Cows entering into the herd during the follow-up were sampled every week only the end of the follow-up (at D7, D14, D21 and D28).

Altogether, 821 individual health states were determined, and respectively 145 complete (i.e. with five sampling points per cow) and 89 incomplete (i.e. with one to four sampling points per cow) temporal trajectories of individual health states were available. These data were here used in the *C. burnetii* model to assess the main epidemic parameters by Bayesian inference. In addition, the same dataset will be used to define and calibrate the representation in the model of the individual variability in shedding routes and levels. This latter point will be described in the section of the thesis dealing with the incorporation in the model of the heterogeneity of shedding (section II of Chapter 3).

Table 2.1. Numbers of animals according to the results of diagnostic tests (ELISA and PCR). The numbers are given for each of the five herds and each of the five sampling points.

Number of animals			**D0**	**D7**	**D14**	**D21**	**D28**
Herd 1	**seronegative**	**PCR negative**	10	6	7	7	5
		PCR positive	2	1	0	0	1
	seropositive	**PCR negative**	7	4	6	6	7
		PCR positive	5	6	4	4	4
	not sampled		0	7	7	7	7
	TOTAL		24	24	24	24	24
Herd 2	**seronegative**	**PCR negative**	23	14	12	16	11
		PCR positive	2	2	3	0	5
	seropositive	**PCR negative**	18	4	9	8	6
		PCR positive	6	7	3	3	5
	not sampled		0	22	22	22	22
	TOTAL		49	49	49	49	49
Herd 3	**seronegative**	**PCR negative**	25	18	17	19	21
		PCR positive	2	1	3	2	0
	seropositive	**PCR negative**	6	5	4	5	4
		PCR positive	1	1	1	1	0
	not sampled		0	9	9	9	12
	TOTAL		34	34	34	36	37
Herd 4	**seronegative**	**PCR negative**	14	14	2	11	21
		PCR positive	7	4	16	10	7
	seropositive	**PCR negative**	4	4	10	6	6
		PCR positive	6	4	0	3	2
	not sampled		0	7	7	7	7
	TOTAL		31	33	35	37	43
Herd 5	**seronegative**	**PCR negative**	28	21	5	17	16
		PCR positive	1	2	10	3	2
	seropositive	**PCR negative**	27	18	33	15	14
		PCR positive	23	12	4	15	19
	not sampled		0	27	28	31	30
	TOTAL		79	80	80	81	81
All herds	**TOTAL**		217	220	222	227	234

2. Description of the data used to define some of the prior distributions (data set B)

Since the estimation was performed in a Bayesian framework (as justified in the next section), we had to define the prior distribution of the health states in an infected dairy cattle herd. To avoid the use of the same data set for both quantification of prior distributions and inference and since additional data were available, we determined the distribution of health states of 251 cows from six French infected dairy cattle herds followed by R. Guatteo (in a different study from the one described in the previous section [61]). The six herds exhibited repeated abortions due to *C. burnetii* confirmed by at least one positive PCR result on vaginal mucus of cow after abortion but no control measure (i.e. antibiotics or vaccination directed against *C. burnetii*) had been implemented before the sampling. The individual state of each cow was consistently determined in the same way as previously described, using an ELISA test on serum and a real-time PCR on milk, faeces and vaginal mucus samples. Table 2.2 shows for each of the six herds the repartition of the cows with respect to their seropositive/seronegative and PCR positive/ PCR negative status. To determine the prior distribution of the health states in an infected dairy cattle herd, we took into account the mean proportions in each of the four categories (see section III of this chapter).

Table 2.2. For each of the six herds, (i) repartition of cows function of their results to diagnostic tests (ELISA and PCR), and (ii) total numbers of cows.

Criteria			Herd 1	Herd 2	Herd 3	Herd 4	Herd 5	Herd 6	Mean
Proportion of cows	Sero -	PCR -	20.9%	46.2%	27.5%	47.4%	50.0%	35.5%	37.9%
		PCR +	14.0%	17.9%	15.7%	8.8%	10.0%	3.2%	11.6%
	Sero +	PCR -	51.2%	10.3%	27.5%	17.5%	20.0%	38.7%	27.5%
		PCR +	14.0%	25.6%	29.4%	26.3%	20.0%	22.6%	23.0%
Total number of cows			43	39	51	57	30	31	251

3. Modelling assumptions

The aim of our study is to understand the spread of *C. burnetii* infection within a dairy herd by assessing the main epidemiological parameters from field data. Since we focused on one population (here the herd), the unit of modelling that we considered was the animal. Based on expert's opinion and observations in data set A, we opted for a modified version of the *SIR*

model: 2 classes of *I* (*I-,* seronegative shedders versus *I+,* seropositive shedders) were considered as important disease categories and represented in the model. Transitions in both directions between *S* and *I-* and between *I+* and *R* were assumed (Figure 2.4). As inhalation of contaminated aerosols is the main route of infection for ruminants, we added a compartment representing the environmental bacterial load and linked the probability of infection (i.e. the probability of transition from *S* to *I-*) to this compartment. As we focused on a population of small size (around 50 cows), all the transitions between health states were supposed stochastic. We chose a time step of a week because no transition could likely occur in less than 7 days. Moreover, only an individual-based model with data at the individual level would allow us to assess the model parameters. In fact, as transitions in both ways between *S* and *I-* and between *I+* and *R* are allowed in the model, the number of animals in each health state at each time point would be an insufficient information. Let us take an example with two health states *A* and *B*. We denote by *p(A=>B)* and *p(B=>A)* the probabilities of transition in each direction for an individual and *NA(t)* and *NB(t)* the total number of individuals in each state at time *t*. Let is say that *NA* and *NB* do not change between two adjacent time points, but that *k* individuals moved in each direction. It is impossible to estimate *p(A=>B)* and *p(B=>A)* if the only information we have are *NA* and *NB*. The inference has to be based on individual trajectories: if we know that *k* individuals moved in each direction, it is perfectly possible to estimate *p(A->B)* as *k/NA* and *p(B->A)* as *k/NB*. This reasoning led us to opt for an individual-based approach which was possible to implement due to the fact that the data we used consisted of individual trajectories (as described in section II.1 of this chapter).

Then, the crucial step was to estimate model parameters from data set A. Several issues were raised: first, the data was incomplete (i.e. some cows were not sampled at each sampling point). Besides, diagnostic tests (and especially ELISA) were imperfect: the individual health state observed in the data could then differ from the real health state of the cow. Lastly, some model parameters were assumed to be herd-dependant. Therefore, we had to deal with the missing data, the uncertainty due to the imperfection of diagnostic tests, and the hierarchical structure of the process to estimate the model parameters. All these arguments converged towards the choice of the Bayesian framework for parameter inference.

III- Manuscript: Spread of Q fever within dairy cattle herds: key parameters inferred using a Bayesian approach

Aurélie Courcoul[1,2], Elisabeta Vergu[3], Jean-Baptiste Denis[3], François Beaudeau[1,2]

[1]INRA, UMR1300 Bio-agression, Epidémiologie et Analyse de Risque, Nantes, France
[2]Oniris, UMR1300 Bio-agression, Epidémiologie et Analyse de Risque, Nantes, France
[3]INRA, UR341 Mathématiques et Informatique Appliquées, Jouy-en-Josas, France

Proceedings of the Royal Society B, 2010 Sep 22; 277(1695):2857-65[2]

1. Abstract

Q fever is a worldwide zoonosis caused by *Coxiella burnetii*. Although ruminants are recognised as the most important source of human infection, no previous studies have focused on assessing the characteristics of the bacterial spread within a cattle herd and no epidemic model has been proposed in this context. We assess the key epidemiological parameters from field data in a Bayesian framework that takes into account the available knowledge, missing data and the uncertainty of the observation process due to the imperfection of diagnostic tests. We propose an original individual-based Markovian model in discrete time describing the evolution of the infection for each animal. Markov chain Monte Carlo methodology is used to estimate parameters of interest from data consisting of individual health states of 217 cows of five chronically infected dairy herds sampled weekly over a four-week period. Outputs are the posterior distributions of the probabilities of transition between health states and of the environmental bacterial load. Our findings show that some herds are characterised by a very low infection risk while others have a mild infection risk and a non-negligible intermittent shedding probability. Moreover, the antibody status seems a key point in the bacterial spread (shedders with antibodies shed for a longer period of time than shedders without antibodies). In addition to the biological insights, these

[2] We gratefully acknowledge the editors of the *Proceedings B* who gave us the permission to reproduce this manuscript in our thesis. The paper is available online: http://rspb.royalsocietypublishing.org/content/277/1695/2857.abstract

estimates also provide information for calibrating simulation models to assess control strategies for *C. burnetii* infection.

2. Introduction

Q fever is a zoonotic disease caused by *Coxiella burnetii,* a bacterium found worldwide in a wide range of animals. Since 2007, Q fever has become an important public health problem in several parts of Europe [72, 108, 121, 141]. Although Q fever in humans is asymptomatic in more than 60% of cases, it may lead to either an acute or a chronic disease [128]. The acute disease is mainly flu-like but severe complications, such as pneumonia or hepatitis can occur. In its chronic form, endocarditis is the most frequent manifestation, especially in patients with pre-existing heart valve lesions. Abortion in pregnant women can also occur. Recently, a large epidemic of Q fever emerged in the southern part of the Netherlands causing more than 3000 human cases since 2007 [33]. A link has been established between some human cases and farms of small ruminants where abortions due to Q fever were detected [141]. Ruminants are recognised as the main source of human infection [109, 118]. Infected animals shed the bacterium through various routes such as parturition products, faeces, urine, vaginal mucus or milk [15, 20, 57]. The transmission of infection both between ruminants and between ruminants and humans is mainly due to inhalation of aerosolised bacteria or contaminated dust [103]. The bacterium survives very well in the environment [167] and can infect humans and animals for a long period after it has been excreted by the host. Therefore, the control of infection within ruminant herds is the most important factor influencing the occurrence of human outbreaks. Besides these obvious implications in terms of public health, controlling the spread of Q fever is also motivated by economic and animal health concerns. Indeed, in ruminants, the infection may also cause abortions, infertility, metritis or chronic mastitis [5, 20, 26, 125].

Previous studies of Q fever in ruminants have shown that some infected animals shed the bacteria in a discontinuous way: this intermittent shedding has been described in the milk and faeces of goats [11] as well as in the milk, faeces and vaginal mucus of cows [37, 59, 131]. However, little information is available on the characteristics of the spread of *C. burnetii* within a cattle herd, a key point in the understanding and the control of the disease. Specifically, the probability that a susceptible cow will become infected when introduced into a chronically infected herd, the duration of shedding for an infectious cow, the differences between the shedding patterns of seronegative and seropositive cows, the probability of intermittent shedding and the duration of non-shedding periods are all key parameters which have not been assessed. In order to address these issues, we propose an

original modelling-based Bayesian approach to quantify the epidemiological parameters related to the transmission of *C. burnetii* within a dairy cattle herd.

We have built a dynamic discrete time individual-based stochastic model describing the evolution of health states with time for each animal. Due to the imperfection of diagnostic tests (assessed by sensitivity (*Se*) and specificity (*Sp*) parameters), the observed health state of a cow in our data can differ from its real health state. Thus, this uncertainty in observations has to be explicitly incorporated in the model to provide more accurate estimates of the parameters, particularly of the transition rates. We use the Bayesian paradigm to deal with this uncertainty, the missing data (since for some animals the health state was not identified at every moment in the follow-up) and to account for the hierarchical structure of the process (e.g. some parameters are herd-dependent). Inference is performed from field data (described in Guatteo et al. [57]) using Markov chain Monte Carlo (MCMC) methodology [53], which is being increasingly used in epidemic modelling [31, 65, 88, 119, 145]. Posterior distributions of model parameters are analysed and biological interpretations are proposed.

3. Data

A one-month longitudinal study was carried out in five French dairy cattle herds infected with *C. burnetii*, but without any clinical sign attributable to Q fever. The selected herds were chosen to satisfy two major criteria: (i) the presence of the bacterium *C. burnetii* within the herd; this was certified by a positive PCR result on bulk tank milk and more than 20% of cows seropositive for *C. burnetii*, and (ii) the absence of any control measure (i.e. antibiotics or vaccination directed against *C. burnetii*) before the end of the study. The protocol of the study is described in detail in Guatteo et al. [57]. To assess the dynamics of *C. burnetii* infection, the lactating cows of these herds were sampled from one to five times on a weekly basis. The observed individual state of each cow was determined at each sampling time using an ELISA test (LSI ELISA Cox Ruminants®, Lissieu, France) on serum and a real-time PCR (LSI Taqvet *Coxiella burneti*®, Lissieu, France) on three different samples (milk, faeces and vaginal mucus). The results of the ELISA test were expressed by the ratio (S/P) between optical densities of the sample and the positive control, and a cow was considered seropositive when the S/P ratio in serum was greater than or equal to 0.4. For the PCR test, only the samples presenting a typical amplification curve (demonstrating *C. burnetii* DNA detection) with a Ct (cycle threshold) below 40 were considered positive. A cow was identified as PCR-positive when at least one of its three samples was PCR-positive. At the initial point of the follow-up (t0), the sizes of the five herds ranged from 24 to 79

lactating cows and a total of 217 cows were tested (see Tables 2.4 and 2.5 of the Supplementary Material). Thereafter, 100% of the initially (at t0) PCR-positive cows, 100% (or 50% in herds with more than 40 lactating cows) of the initially seropositive/PCR-negative cows, and 65% of the initially seronegative/PCR-negative cows were retained for the follow-up. Thus, during the following month, between 55% and 79% of the cows of each herd were tested every week (at t7, t14, t21 and t28) in the same way in order to determine their individual health state. The cows entering one of the herds during the study (as a consequence of a purchase or a first calving) were also included. According to the PCR results and the ELISA test, at t0 between 35% and 74% of cows per herd were identified as PCR-negative/seronegative, between 1% and 23% were PCR-positive/seronegative, between 2% and 35% were PCR-positive/seropositive and between 17% and 37% were PCR-negative/seropositive. At the end point of the follow-up (day 28 – t28), the herds comprised between 24 and 81 lactating cows. Altogether, 821 individual health states were determined and 235 (complete or incomplete) temporal trajectories of individual health status were available.

4. Model and methods

Based on the available knowledge concerning the clinical and epidemiological aspects of Q fever, an epidemic model describing its spread within a dairy cattle herd was built. Firstly, the allowed transitions between the health states of the epidemiological model are described. Then, the dynamic model representing the temporal evolution of observed individual health states is presented. Finally, we detail the assumed priors and calculated posterior distributions of the model parameters in the Bayesian framework (using MCMC methods).

a. Epidemic model

Each individual of the population of lactating cows is in one of four mutually exclusive health states at a given time, as shown in Figure 2.4. By inhaling bacteria contained in the environment, a susceptible cow, *S* (non-shedder without antibodies), can become infectious, *I-* (shedder without antibodies), and start shedding. Either it manages to eliminate the bacterium and becomes *S* again (non-shedder without antibodies and then apparently susceptible) or it produces antibodies and continues being infectious and shedding, *I+ (*shedder with antibodies). When it stops shedding, it becomes *R* (non-shedder with antibodies). Since the shedding is intermittent [59, 131], a transition from *R* to *I+* is assumed. Antibodies can last several years in humans [47] and at least several months in

cattle [125]. Here, we assume that the probability of observing a cow lose its antibodies over the period of study (one month) is very low and negligible, especially in chronically infected herds where immunity is probably steadily stimulated. Therefore, no transition from health states with antibodies (*I+* or *R*) to health states without antibodies (*I- or S)* is allowed in our model. Shedders (*I-* and *I+*) contribute to filling the environment compartment (*E*) with the bacteria: ε_1 and ε_2 are the quantities of bacteria shed during a time step (one week in our case) by an individual *I-* and *I+* respectively. The probability of infection or re-infection, *p* (transition from *S* to *I-*) is expressed at each time step as $p_t = 1-\exp(-E_t)$, where E_t is the quantity of bacteria in the environment of the herd at time *t* (one unit of E_t corresponding to a probability of transition from *S* to *I-* of (1-1/e)). The mortality rate of *C. burnetii* in the environment, μ includes the natural mortality of the bacterium and its removal in relation to the periodic cleaning of the cattle housing carried out by the farmer.

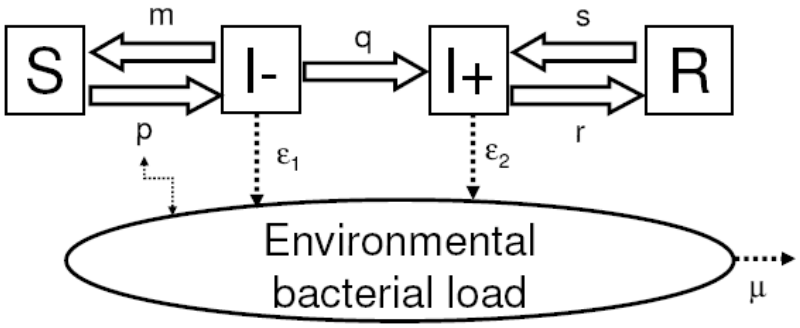

Figure 2.4. Flow diagram describing the modelled spread of *C. burnetii* within a cattle herd. The health states are: *S,* non-shedder cow without antibodies, *I-*, shedder cow without any antibodies, *I+*, shedder cow with antibodies and *R*, non-shedder cow with antibodies. *E* represents the environmental bacterial load. The model parameters are: *p*, the probability of infection or reinfection (equal to $1-\exp^{(-E)}$), *m*, the probability of transition from *I-* to S, *q*, the probability of transition from *I-* to *I+*, *r,* the probability of transition from *I+* to *R*, *s*, the probability of transition from *R* to *I+*, ε_1 and ε_2, the quantities of bacteria shed during a time step by an individual *I-* and *I+* respectively and μ, the mortality rate of *C. burnetii* in the environment.

b. Bayesian framework

We develop a dynamic discrete time individual-based stochastic model to represent the temporal evolution of the observed health state of each cow. This is done in two main steps: firstly, the temporal evolution of the real individual health state is modelled using Markovian transitions and secondly, the uncertainty of the observations is incorporated using the *Se* and *Sp* of the two diagnostic tests.

Let $R_{t,h}^{(i)} \in \{S, I-, I+, R\}$ be the real health state of individual i belonging to herd h ($i \in \{1,\ldots,N(h)\}$ with $N(h)$ the total number of cows in the herd h, $h \in \{1,\ldots,H\}$ and H the number of herds) at time t ($t \in \{0,\ldots,T\}$ with t28=T and t0=0). As illustrated by the graph in Figure 2.5, for t>t0, $R_{t,h}^{(i)}$, depends on $R_{t-1,h}^{(i)}$ and on $E_{t,h}$, the quantity of bacteria in the environment of herd h at time t. The transition probabilities can be gathered in the matrix $Q_{t,h}$:

$$Q_{t,h} = \begin{pmatrix} 1-p_{t,h} & p_{t,h} & 0 & 0 \\ m & 1-m-q & q & 0 \\ 0 & 0 & 1-r & r \\ 0 & 0 & s_h & 1-s_h \end{pmatrix}, \text{ where } Q_{t,h,jk} = P(R_{t,h}^{(i)} = x_k \,\middle|\, R_{t-1,h}^{(i)} = x_j) \quad (3.1)$$

for $t=1,\ldots,T$, $i=1,\ldots,4$ and $x_j, x_k \in \{x_1= S, x_2=I-, x_3=I+, x_4=R\}$.

The transition probability from S to $I-$ varies with time and herd since $p_t = 1-\exp(-E_t)$. This is not the case for the other transition probabilities: m, q and r are assumed constant. As s is related to the intermittency of shedding, possibly due to a stress specifically occurring in a given herd (like an anti-parasitic treatment or a modification in herd management), this parameter is considered herd-dependent.

The initial real health states, $R_{0,h}^{(i)}$, are independent random variables with a probability distribution specified by J, where $J_{x_j} = P(R_{0,h}^{(i)} = x_j)$ for $x_j \in \{x_1= S, x_2=I-, x_3=I+, x_4=R\}$. The environment dynamics is expressed by the equation: $E_{t+1,h} = (1-\mu)E_{t,h} + \varepsilon_1 I_{t,h}^{-} + \varepsilon_2 I_{t,h}^{+}$, as it is dependent on the quantity of bacteria in the environment and the prevalence of shedders ($I_{t,h}^{-}$, $I_{t,h}^{+}$) at the previous time (Figure 2.5). Since the beginning of the follow-up does not correspond to the infection onset, the initial content of *C. burnetii* in the environment of each herd, $E_{0,h}$, is not zero and has to be introduced and then estimated.

The observation level accounts for the uncertainty of the observations $O_{t,h}^{(i)}$ and describes their relationship with the real health states $R_{t,h}^{(i)}$ using the matrix U:

$$U = \begin{pmatrix} Sp_{PCR}Sp_{EI} & (1-Sp_{PCR})Sp_{EI} & (1-Sp_{PCR})(1-Sp_{EI}) & Sp_{PCR}(1-Sp_{EI}) \\ (1-Se_{PCR})Sp_{EI} & Se_{PCR}Sp_{EI} & Se_{PCR}(1-Sp_{EI}) & (1-Se_{PCR})(1-Sp_{EI}) \\ (1-Se_{PCR})(1-Se_{EI}) & Se_{PCR}(1-Se_{EI}) & Se_{PCR}Se_{EI} & (1-Se_{PCR})Se_{EI} \\ Sp_{PCR}(1-Se_{EI}) & (1-Sp_{PCR})(1-Se_{EI}) & (1-Sp_{PCR})Se_{EI} & Sp_{PCR}Se_{EI} \end{pmatrix} \quad (3.2)$$

where $U_{jk} = P(O_{t,h}^{(i)} = x_k \mid R_{t,h}^{(i)} = x_j)$, for $t=0..T$, $i=1..4$ and x_i, $x_k \in \{x_1= S, x_2=I-, x_3=I+, x_4=R\}$.

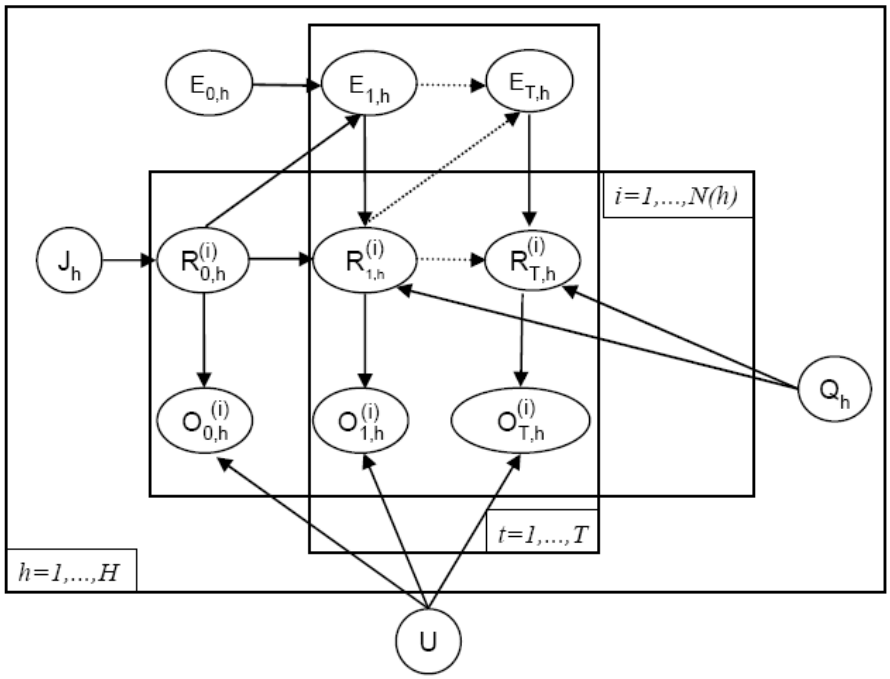

Figure 2.5. Network describing the temporal evolution of individual health states of animals within an infected dairy cattle herd. $R_{t,h}^{(i)} \in \{S, I-, I+, R\}$ represents the real and non-observed health state of individual *i* belonging to herd *h* (i ∈ {1,…N(h)} with *N(h)* the total number of cows in the herd *h*, *h* ∈ {1,…H} and *H* the number of herds) at time *t* (t ∈ {0,…T)} with t28=T and t0=0). $E_{t,h}$ describes the quantity of bacteria in the environment of the herd *h* at time *t*. $O_{t,h}^{(i)}$ represents the observed health state associated with $R_{t,h}^{(i)}$. $\mathbf{J_h}$ is the probability distribution of the initial real health states and **U** is the matrix of the uncertainty parameters (*Se* and *Sp* of tests) linking real and observed health states. $\mathbf{Q_h}$ contains the parameters of transitions between real health states in herd *h* except those characterising the *S*<->*I-* transitions. $\mathbf{Q_h}$ is a 3x4 matrix corresponding to the last three rows of matrix $\mathbf{Q_{t,h}}$ described in Equation (3.1).

We consider that the assumption of conditional independence between ELISA and PCR is reasonable because the two tests have different bases: ELISA relies on the detection of antibodies while PCR is a DNA-based technique to detect bacteria. Enoe et al. [41] made the same assumption to assess the sensitivities and specificities of a nested PCR and a microscopic examination of kidney imprints for the detection of *Nucleospora salmonis* in rainbow trout. Elements of U are then defined as combinations of the specificities of the

PCR and ELISA tests (Sp_{PCR} and Sp_{El} respectively) and their respective sensitivities (Se_{PCR} and Se_{El}).

c. *Bayesian inference: calculation of the posterior distribution of the model parameters from likelihood and prior distribution*

In the Bayesian paradigm, the joint posterior distributions of model parameters can be written as $p(J,Q|O) \propto L(O|J,Q) * \pi(J,Q)$, where $L(O|J,Q)$ *and* $\pi(J,Q)$ are the likelihood function and the joint prior distribution of model parameters respectively and $Q = \bigcup_{\substack{t=1...T \\ h=1...5}} Q_{t,h}$ (see subsection 8. Supplementary Material for more details).

Since the uncertainty parameters of the matrix *U* are fixed, they are not considered in the joint prior density $\pi(J,Q)$. The *Se* of the ELISA test is set equal to 0.85 (according to a recent estimation, Guatteo, *pers. comm.)* and the *Sp* is taken as equal to 0.95, while for the real-time PCR, both *Se* and *Sp* are fixed at 0.95. As no published data on the test characteristics are available, these values were chosen in accordance with expert opinion.

Available knowledge is incorporated into the model through prior distributions. Given that *C. burnetii* withstands hard environmental conditions [103], the median of its life expectancy ($1/\mu$) on the farm in an infectious form is considered to be 4.5 weeks with a 95% credible interval (CI) of 0.7-14 weeks. To determine the prior distribution of the initial real health state *J*, we use independent data from six other French infected dairy cattle herds. On average per herd 38% (min=20.9%, max=50%) of cows were observed to be in state *S*, 12% (3.2%, 17.9%) in state *I-*, 27% (10.3%, 51.2%) in state *I+* and 23% (14%, 29.4%) in state *R* (Guatteo, *pers. comm.)*. As the initial proportions of *S, I-, I+* and *R* should sum to one (as they represent a partition of the individual health states), an appropriate prior distribution of the initial health state *J* is a Dirichlet distribution, D (3.5, 1, 2.5, 2). Its coefficients are chosen to account for the observed proportions in the extra data (e.g. proportion of *S* is 3.5/9=38%, etc). Concerning the transition parameters (*p*, *m*, *q*, *r* and *s*), minimally informative prior densities, reflecting the lack of information, are chosen. As these parameters are assumed to lie between 0 and 1, Beta distributions are used for the probabilities of transition from *S* to *I-* (*p*), *I+* to *R* (*r*) and *R* to *I+* (*s*) health states. A Dirichlet distribution is assumed for the probabilities of transition from *I-* to *S (m)* and *I+ (q),* respectively, since the sum of *m, q and "the probability of staying in I-"* is equal to 1. The marginal distributions of *m*, *q*, and Beta distributions for *r* and *s* are rather flat. As *C.*

burnetii spreads moderately quickly in cows [8, 66], we use a Beta distribution with a median of 0.33 and a 95% CI=[0.05-0.77] in order to penalise high values of p. As the environmental bacterial load E can be expressed with respect to the probability of infection p, the prior on E is deduced from the prior on p (median=0.4, 95% CI=[0.05-1.44]). Concerning the excretion parameters ε, we faced a complete lack of information. However, as ε is the quantity of bacteria shed per week by a shedder cow, a plausible assumption is that $\varepsilon=\varepsilon_1+\varepsilon_2$ is lower than the environmental bacterial load E. Hence, we use a truncated Normal distribution with a median of 0.23 and 95% CI=[0-0.72] for both ε_1 and ε_2. All these prior distributions are detailed in Table 2.3.

Since posterior distributions are not analytically tractable, inference is based on computationally intensive methodology: MCMC methods based on the Gibbs sampling algorithm implemented in JAGS 1.0.3 are used. Bayesian MCMC allows datasets with missing data to be fully modelled by sampling missing data points from the posterior distributions (in Equation S1 of the Supplementary Material the matrix O is not entirely observed). Results are analysed with R 2.8.1 [127] and R package coda [126].

d. Model adequacy

In order to check the model adequacy for the data, a subsequent assessment is performed. We simulate infection spread in five cattle herds with the same size, same initial environmental content and same number of missing data as in the original dataset, using parameters drawn from inferred posterior distributions. The missing pattern (i.e. missing data during the follow-up are more frequent for PCR-negative cows at t0 than for PCR-positive ones) is not taken into account. The quantiles of the numbers of transitions between observed health states in each herd for a time interval of one week are calculated and compared with the data.

5. Results

Visual inspection of the chain pattern does not indicate non-convergence of the MCMC algorithm (results not shown). Most of the parameters have a potential scale reduction factor of the Gelman-Rubin diagnostic [51] close to one (≤1.05). However, five of the 35 independent parameters monitored have values of potential scale reduction factors between 1.05 and 1.27. For these parameters, the results have to be interpreted with care (see Table 2.6 of the Supplementary Material for details). Median values and 95% CI of posterior densities (represented in Figures 2.7 and 2.8 of subsection 8. Supplementary Material) of inferred parameters are given in Table 2.3 and in subsection 8 (Table 2.7).

Table 2.3. Priors and posteriors for the model parameters. For the posterior distributions, medians and 95% credible intervals (CI) are shown.

Model parameter	Prior distribution or mathematical expression	Prior median and 95% CI	Posterior median and 95% CI				
			Herd 1	Herd 2	Herd 3	Herd 4	Herd 5
Transition rate I- -> S, m (week^{-1})	Dirichlet (α_1, α_2, α_3), α_i~Uniform(0.5,10)	0.32 (0.01 - 0.76)	0.695(0.542 – 0.844)				
Transition rate I- ->I+, q (week^{-1})			0.017 (0.001 - 0.082)				
Time spent in state I- by an individual (weeks)	$1/(m+q)$	1.4 (1.0 - 4)	1.4 (1.1 - 1.8)				
Transition rate I+ -> R, r (week^{-1})	Beta(1,1)	0.50 (0.03 - 0.98)	0.204 (0.121 - 0.294)				
Time spent in state I+ by an individual (weeks)	$1/r$	2.0 (1.0 - 33.3)	4.9 (3.4 - 8.3)				
Transition rate R -> I+, s (week^{-1})	Beta(1,1)	0.50 (0.02 - 0.98)	0.037 (0.006 - 0.159)	0.170 (0.042 - 0.345)	0.036 (0.006 - 0.162)	0.380 (0.070 - 0.737)	0.277 (0.109 - 0.466)
Time spent in state R by an individual (weeks)	$1/s$	2.0 (1.0 – 34.1)	26.6 (6.3 - 159.9)	5.9 (2.8 - 23.9)	28.1 (6.1 - 163.7)	2.6 (1.3 - 14.3)	3.6 (2.1 - 9.2)
Quantity of bacteria shed by an I- per week, ε_1	Normal(0.25,0.25)	0.30 (0.02 - 0.76)	0.003 (0.000 - 0.018)				
Quantity of bacteriashed by an I+ per week, ε_2	Normal(0.25,0.25)	0.30 (0.02 - 0.76)	0.004 (0.000 - 0.016)				
Mortality rate of the bacterium, μ (week^{-1})	$1/\mu$ ~ Beta(2.16,38.51)	0.21 (0.07- 0.80)	0.289 (0.079 - 0.926)	0.274 (0.079 - 0.869)	0.289 (0.082 - 0.874)	0.171 (0.068 - 0.407)	0.203 (0.071 -0.760)
Transition rate S -> I-, p_1 (infection risk at t0)	Beta(1.77,3.3)	0.33 (0.05 - 0.77)	0.135 (0.021 -0.436)	0.124 (0.024 -0.325)	0.073 (0.014 - 0.213)	0.466 (0.272 - 0.660)	0.230 (0.044 - 0.455)
Bacterial load in the environment at t0, E_0	$-\log(1-p_1)$	0.40 (0.05 - 1.46)	0.146 (0.021 -0.572)	0.133 (0.024 -0.392)	0.076 (0.014 - 0.240)	0.627 (0.318 - 1.078)	0.261 (0.045 - 0.606)
Transition rate S -> I-, p_5 (infection risk at t28)	$1-\exp^{(-E4)}$	variable function of herds	0.095 (0.010 -0.282)	0.136 (0.024 - 0.334)	0.044 (0.005 - 0.134)	0.403 (0.204 -0.622)	0.428 (0.182 - 0.722)
Bacterial load in the environment at t28, E_4	$E_4 = (1-\mu)E_3+\varepsilon_1$*I- +$\varepsilon_2$*I+	variable function of herds	0.100 (0.010 - 0.331)	0.146 (0.024 - 0.406)	0.044 (0.005 - 0.143)	0.517 (0.228 - 0.973)	0.558 (0.201 - 1.278)

a. Parameters of transition between health states

The spread of *C. burnetii* within a dairy cattle herd is mainly characterised by shedding parameters and probabilities of transition between health states, which are also interpretable as sojourn times in these states (equal to the inverse of transition parameters). For all these parameters, the posterior distributions cover shorter intervals than those defined by the prior distributions, which reveals that the data provide information. The probability of transition from the non-infection state *S* to the shedder state *I-* (corresponding to the infection risk) seems moderate in some herds (for example in herd 3 with a median *p* at time 1 of 0.073 and a 95% CI=[0.014–0.213]) but quite high in others (for example in herd 4 with a median *p* at t0 of 0.466 and 95% CI=[0.272-0.660]). Whereas the transition from the shedder state without antibodies, *I-*, to the non-infected state *S* is relatively more rapid (median of *m* equal to 0.695 $week^{-1}$, 95% CI=[0.542-0.844]), the acquisition of antibodies in the infectious state (transition *I- -> I+*) is rather rare (median of *q* equal to 0.017 $week^{-1}$, 95% CI =[0.001-0.082]. Moreover, the duration in health state *I-* is shorter than in *I+* : posterior distributions do not overlap and if we compare the medians, the median duration in *I-* is more than three times shorter than that in *I+* (1.4 versus 4.9 weeks respectively). The median time spent in state *R* before new shedding (representing the intermittency of shedding) is less than 3.6 months in two of the five herds (herds 4 and 5) but can potentially be longer in the other three (e.g. 26.6, 95% CI=[6.3-159.9] in herd 1).

b. Environment-related parameters

Concerning the shedding parameters, as the posterior distributions of the quantities of bacteria excreted by infectious cows without antibodies (ε_1) and with antibodies (ε_2) are almost superimposed, we cannot determine if *I-* animals shed more than, at a similar level to, or less than *I+* animals. For all but herd 5, the posterior distributions of the environmental bacterial load do not vary much with respect to time (Figure 2.8 of subsection 8). Therefore, it is not possible to know how the environmental bacterial load evolves with time. For herd 5, as the posterior distribution shifts to the right from t0 to t28 it is possible that the environmental bacterial load increases with time (at t0: median of 0.261, 95% CI=[0.045-0.606], at t28: median of 0.558, 95% CI=[0.201-1.278]). Since at a given time posterior distributions of *E* widely overlap, we can not determine if environmental bacterial loads differ between herds. For the parameter μ, the posterior distributions are close to the prior distribution regardless of the herd. It seems that the dataset does not contain sufficient information to assess this parameter.

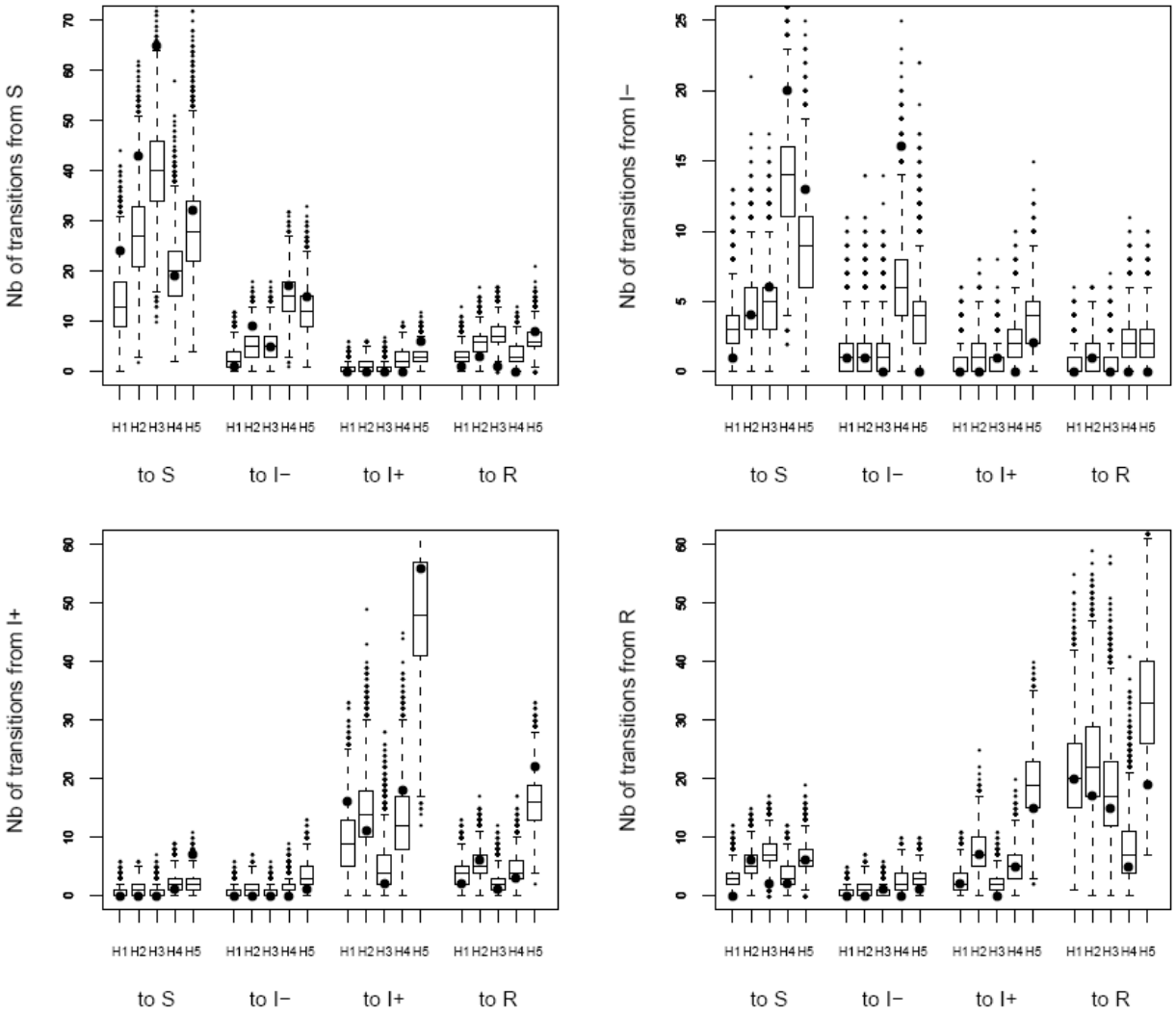

Figure 2.6. Goodness-of-fit assessment. Boxplots summarise the posterior predictive distributions of simulated numbers of weekly transitions between observed health states (*S*, *I-*, *I+* and *R)* in each herd (H1 to H5) during the one-month follow-up. The quartiles are represented by horizontal lines. The whiskers indicate maximum and minimum values of the simulated distributions that lie less than 1.5 IQR lower or higher than the first or the third quartiles respectively. Simulated values beyond the ends of the whiskers are indicated by a point. Dark filled-in circles represent numbers of transitions between observed health states in our dataset.

c. Checking of model adequacy for the data

The goodness-of-fit is assessed in Figure 2.6. We verify the ability of the model to reproduce observed summary statistics, defined as the total number of transitions per week between observed individual health states for each herd, during a month, when parameters are sampled from posterior densities. Sixty-three percent (expected 50%) of observed summary statistics lie within the predicted 50% CI and 94% (expected 95%) of them belong to the 95% CI of the simulated numbers of transitions.

6. Discussion

This study, based on a Bayesian modelling approach, provides the first quantitative assessment of parameters describing the spread of *C. burnetii* within chronically infected dairy herds. Previous studies that focused on Bayesian statistical inference of disease parameters have already proposed discrete time stochastic epidemic models [88, 114]. However, our approach differs from these as it is individual-based.

The Bayesian framework enables the combination in the same model of previous knowledge about *C. burnetii* (mainly concerning the life expectancy of the bacteria in the environment and the proportions of different health states within an infected herd) with information coming from the present dataset. Moreover, it allows differences between herds to be accounted for in a flexible manner through a hierarchical representation of the processes involved. The convergence of the MCMC is not perfect, particularly for the initial real health states. Although estimations of these parameters seem biologically consistent, our dataset is probably not informative enough to provide good assessments of all inferred parameters. However, for most of the parameters, convergence is achieved, the results are biologically plausible and the goodness-of-fit is satisfactory overall. Nevertheless, the choice of simulated missing data (that is, of cows with unknown health states for the t7-t28 period) was made randomly whereas in the field protocol, the selection of the weekly sampled cows was not made at random. Moreover, a possible way to improve further the adequacy of the model for the data is to consider that the uncertainty on the observed health states would differ for each observation as a function of the quantitative results provided by the diagnostic tests (S/P ratios for the ELISA and Ct values for the PCR). In fact, dichotomising the test result of an ELISA can be unnecessary and, to some extent, counter-productive [116]. The relevancy of this option could be explored in further studies.

As shown by the present results, some chronically infected herds (like herd 3) are characterised by a low probability of infection and then a slow spread of the disease while

others (like herd 4) are characterised by a quite high probability of infection and then a faster infection dynamics. Also, intermittency of shedding is less likely to occur in some herds (like herds 1 and 3) but seems usual in others (like herd 5).

When a cow becomes infected, clearance of the bacterium without seroconversion (transitions from $I-$ to S) is very common while the transition from the seronegative to the seropositive state ($I- \rightarrow I+$) is very rare, which means that very few cows of the analysed dataset seroconverted over the month studied (which did not correspond to the beginning of the infection). Moreover, in herds where the infection dynamics is faster, some cows are restrained to transitions between the non-infected state and shedding without antibodies state ($S \leftrightarrow I-$), while others are restrained to transitions between the infectious seropositive state and the non-shedding seropositive state ($I+ \leftrightarrow R$). Thus, two categories of animals seem to exist with two different types of infection response: a response with or without any antibody production. Lastly, the antibody status seems to play a major role in the involvement of a given cow in the bacterial spread: shedders with antibodies ($I+$) release bacteria for a longer time than animals in the shedding state without antibodies.

Estimations of the environmental bacterial load are also provided. Although these values do not have any obvious biological meaning, they are related to the infection/re-infection probability of an animal within an infected herd. Our results do not show if the infection risk varies with time but it is likely that some herds (like herd 5 and maybe herd 4 at the end of the study) have quite high infection risks. As the present dataset does not contain enough information to update significantly the prior distribution of the mortality rate of *C. burnetii* (parameter μ), we cannot claim that this potential high probability of infection is due to an ineffective cleaning process of the cattle housing or is directly related to differences in the prevalence of shedding cows. Further work is needed to provide relevant indicators of the environmental contamination. The time scale of our study is probably insufficient to investigate environmental content variations; a period longer than one month is likely required.

The present data do not distinguish real susceptible individuals from non-shedding seronegative ones: all are gathered in the unique category S. Thus, the estimated transition rate from the non-shedding to the shedding without antibodies state is a mix between an infection rate and a re-infection rate. These two rates are different as, in the latter, the cell immunity should already have been activated. However, it is not possible with current diagnostic tests to differentiate primary infected from re-infected animals. The relevance

of cell immunity tests (i.e. skin tests) to study the immunity responses in chronically infected herds would be a profitable area of research.

To conclude, this work provides the first quantitative estimation of key parameters from field data based on an original modelling approach, enabling a better understanding of *C. burnetii* infection dynamics within chronically infected dairy herds. Besides the biological insights provided by the estimated values of parameters, the outputs can be further used to calibrate a simulation model representing the infection dynamics within a cattle herd over a longer time scale and assessing the effectiveness of different control strategies for *C. burnetii* infection.

7. Acknowledgements

The authors thank Annie Rodolakis and Raphaël Guatteo for useful discussions on data and biological aspects related to Q fever and the referees for their comments and suggestions.

8. Supplementary material

a. Data

A one-month longitudinal study was carried out in five French dairy cattle herds infected with *Coxiella burnetii*. Table 2.4 provides the number of the individual health states of the lactating cows of each herd at the beginning (t0) and at the end (t28) of the follow-up.

Table 2.4. Description of the five studied herds at t0 and t28 (aggregated data).

Herd	Cow status at t0					Cow status at t28					
	S	I-	I+	R	Total	S	I-	I+	R	Unknown	Total
1	10	2	5	7	24	5	1	4	7	7	24
2	23	2	6	18	49	11	5	5	6	22	49
3	25	2	1	6	34	21	0	0	4	12	37
4	14	7	4	6	31	21	7	6	2	7	43
5	28	1	27	23	79	16	2	14	19	30	81

Examples of individual trajectories for some cows of herd 2 are given below. See paragraph 4.a. of this section ('Model and methods' – 'Epidemic model') for the definition of the different health states *S*, *I-*, *I+* and *R*.

Table 2.5. Evolution of observed individual health states over time for some cows of the data set.

Cow number	t0	t7	t14	t21	t28
214	S	S	S	S	I-
218	S	S	S	S	S
220	I+	I+	R	I+	I+
222	R	I+	R	R	R
224	R	R	R	R	R
233	I+	I+	I+	I+	I+
234	R	I+	R	R	R
235	S	S	S	S	S
239	I-	I-	S	S	S
240	I+	I+	I+	I+	I+

b. Likelihood

From Equations (3.1) and (3.2) of the main text and considering that all variables are categorically distributed (i.e. they follow a multinomial distribution with the parameter n fixed at 1), the likelihood function of the complete data is given by:

$$L(O|J,Q)=\prod_{\substack{h=1,\ldots,H\\ i=1,\ldots,N(h)}}\left(\sum_{R^i_{0,h},\ldots,R^i_{T,h}} J_{R^{(i)}_{0,h}}\prod_{t=0}^{T}U_{O^{(i)}_{t,h}R^{(i)}_{t,h}}\prod_{t=1}^{T}Q_{R^{(i)}_{t,h}R^{(i)}_{t-1,h}}\right), \qquad \text{(S1)}$$

where random variables are assimilated to their realisations when used as indexes for reasons of simplicity (e.g. $U_{O^{(i)}_{t,h},R^{(i)}_{t,h}} = P(O^{(i)}_{t,h}=x_k \mid R^{(i)}_{t,h}=x_j) = U_{x_j x_k} = U_{jk}$).

c. Convergence of the MCMC algorithm

Three chains were run: an initial burn-in of 10,000 runs with a thin interval of 600 was performed. Then, 50,000 iterations with the same thin interval were run. This thin interval ensures that the chains are no longer autocorrelated. All the 50,000 iterations were used to assess the posterior distributions. A total of 60 parameters were monitored: 25 for $E_{t,h}$, the environmental bacterial load at every sampling time in each herd, 20 for J_h, the distribution of the initial real health states in each herd, five for μ, the herd-dependent mortality rate of *C. burnetii*, five for s, the herd-dependent transition rate for R to $I+$ and one for ε_1, ε_2, m, q and r, the shedding and transition parameters. Among these 60 parameters, 35 are independent: 15 for J_h, five for μ,, five for s, five for p at time 1 of the follow-up and one for ε_1, ε_1, m, q and r.

Moreover, Table 2.6 provides the Gelman-Rubin convergence diagnostic (or the potential scale reduction factor) for these 35 parameters.

The JAGS code used to make Bayesian inference is available on request from A. Courcoul (aurelie.courcoul@oniris-nantes.fr).

Table 2.6. Median and 97.5% percentile of the Gelman-Rubin potential scale reduction factors (PSRF) for the 35 independent parameters of the model. The multivariate PSRF is equal to 1.51.

Parameter	Median PSRF	97.5% percentile of the PSRF
p_1: transition rate S => I- time 0, herd 1	1.03	1.03
p_2: transition rate S => I- time 0, herd 2	1.01	1.02
p_3: transition rate S => I- time 0, herd 3	1.00	1.01
p_4: transition rate S => I- time 0, herd 4	1.00	1.01
p_5: transition rate S => I- time 0, herd 5	1.01	1.02
m: transition rate I- => S	1.01	1.02
q: transition rate I- => I+	1.00	1.01
r: transition rate I+ => R	1.01	1.02
s_1: transition rate R => I+ - herd 1	1.04	1.13
s_2: transition rate R => I+ - herd 2	1.08	1.23
s_3: transition rate R => I+ - herd 3	1.00	1.02
s_4: transition rate R => I+ - herd 4	1.01	1.03
s_5: transition rate R => I+ - herd 5	1.14	1.42
ε_1: quantity of bacteria shed by an I- per week	1.01	1.03
ε_2: quantity of bacteria shed by an I+ per week	1.08	1.24
μ_1: mortality rate of the bacterium, herd 1	1.03	1.10
μ_2: mortality rate of the bacterium, herd 2	1.04	1.13
μ_3: mortality rate of the bacterium, herd 3	1.00	1.00
μ_4: mortality rate of the bacterium, herd 4	1.00	1.01
μ_5 mortality rate of the bacterium, herd 5	1.01	1.02
$J_{1,1}$: proportion of S as initial real health state, herd 1	1.10	1.31
$J_{2,1}$: proportion of S as initial real health state, herd 2	1.05	1.17
$J_{3,1}$: proportion of S as initial real health state, herd 3	1.01	1.03
$J_{4,1}$: proportion of S as initial real health state, herd 4	1.00	1.01
$J_{5,1}$: proportion of S as initial real health state, herd 5	1.27	1.74
$J_{1,2}$: proportion of I- as initial real health state, herd 1	1.00	1.00
$J_{2,2}$: proportion of I- as initial real health state, herd 2	1.00	1.00
$J_{3,2}$: proportion of I- as initial real health state, herd 3	1.02	1.08
$J_{4,2}$: proportion of I- as initial real health state, herd 4	1.00	1.01
$J_{5,2}$: proportion of I- as initial real health state, herd 5	1.02	1.08
$J_{1,3}$: proportion of I+ as initial real health state, herd 1	1.00	1.00
$J_{2,3}$: proportion of I+ as initial real health state, herd 2	1.00	1.00
$J_{3,3}$: proportion of I+ as initial real health state, herd 3	1.03	1.10
$J_{4,3}$: proportion of I+ as initial real health state, herd 4	1.00	1.01
$J_{5,3}$: proportion of I+ as initial real health state, herd 5	1.01	1.03

d. Posterior and prior distributions of transition and shedding parameters

Figure 2.7 provides the prior and posterior distributions of four of the five transition rates between health states (*p*, *m*, *q*, *r* and *s*) and of the two shedding parameters ε_1 and ε_2. As *p*, the transition rate from *S* to *I-* (which represents the infection risk) and *s*, the transition rate from *R* to *I+* (which represents the intermittent shedding) are assumed to be herd-specific, five posterior distributions (one per herd) are given for these two parameters. For all these parameters, the posterior distributions cover shorter intervals than those defined by the prior distributions, which reveals that the data provide some information.

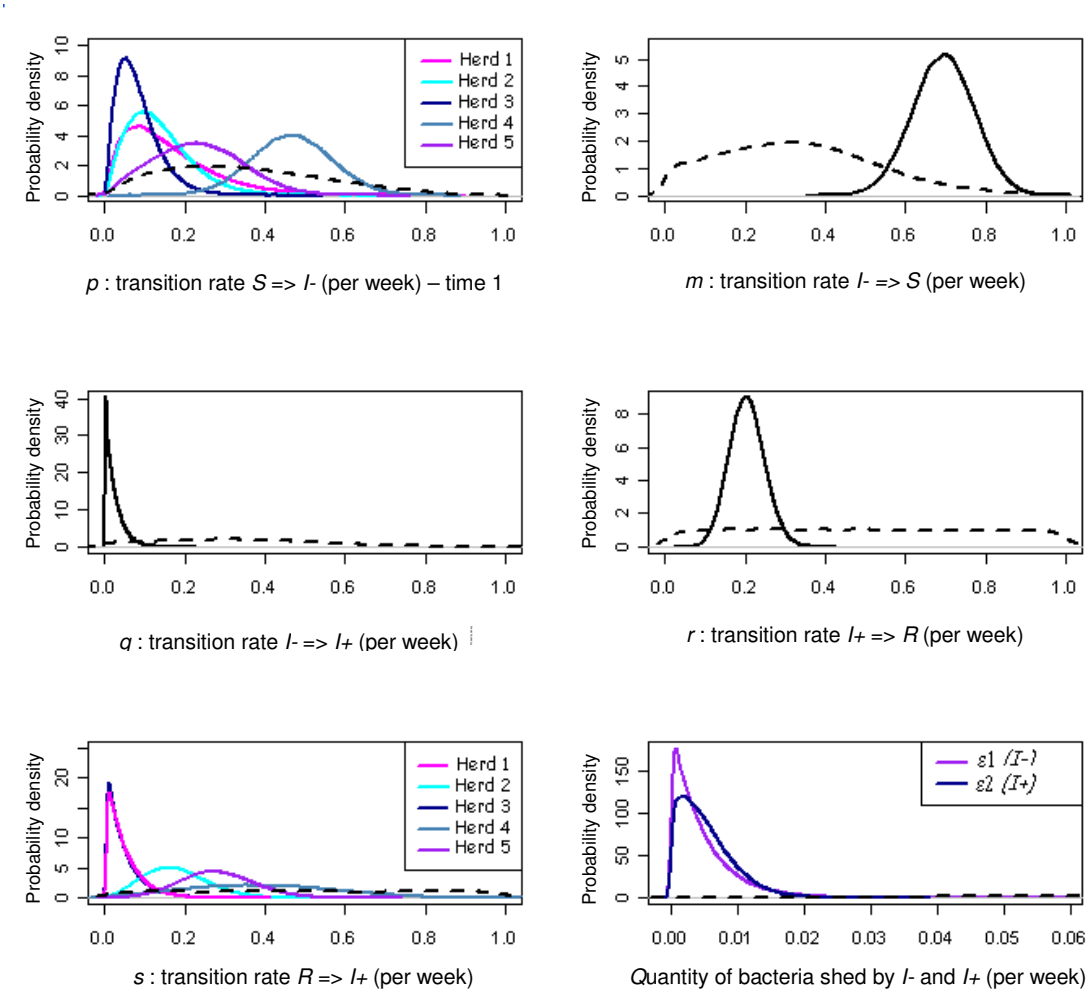

Figure 2.7. Prior (dotted black line) and posterior (solid lines) distributions of the model transition parameters: transition rate from *S* to *I-* (*p*), transition rate from *I-* to *S* (*m*), transition rate from *I-* to *I+* (*q*), transition rate from *I+* to *R* (*r*), transition rate from *R* to *I+* (*s*), quantity of bacteria shed by an *I-* individual in a week (ε_1) and quantity of bacteria shed by an *I+* individual in a week (ε_2).

e. Posteriors and priors of the environment

Figure 2.8 provides the posterior distributions for the environmental bacterial load (E) for each herd at each time and for μ, the mortality rate of the bacterium (which comprises the natural mortality of the bacterium and its removal in relation to the periodic cleaning of the cattle housing carried out by the farmer). For the initial environmental bacterial load and for μ, priors are also given.

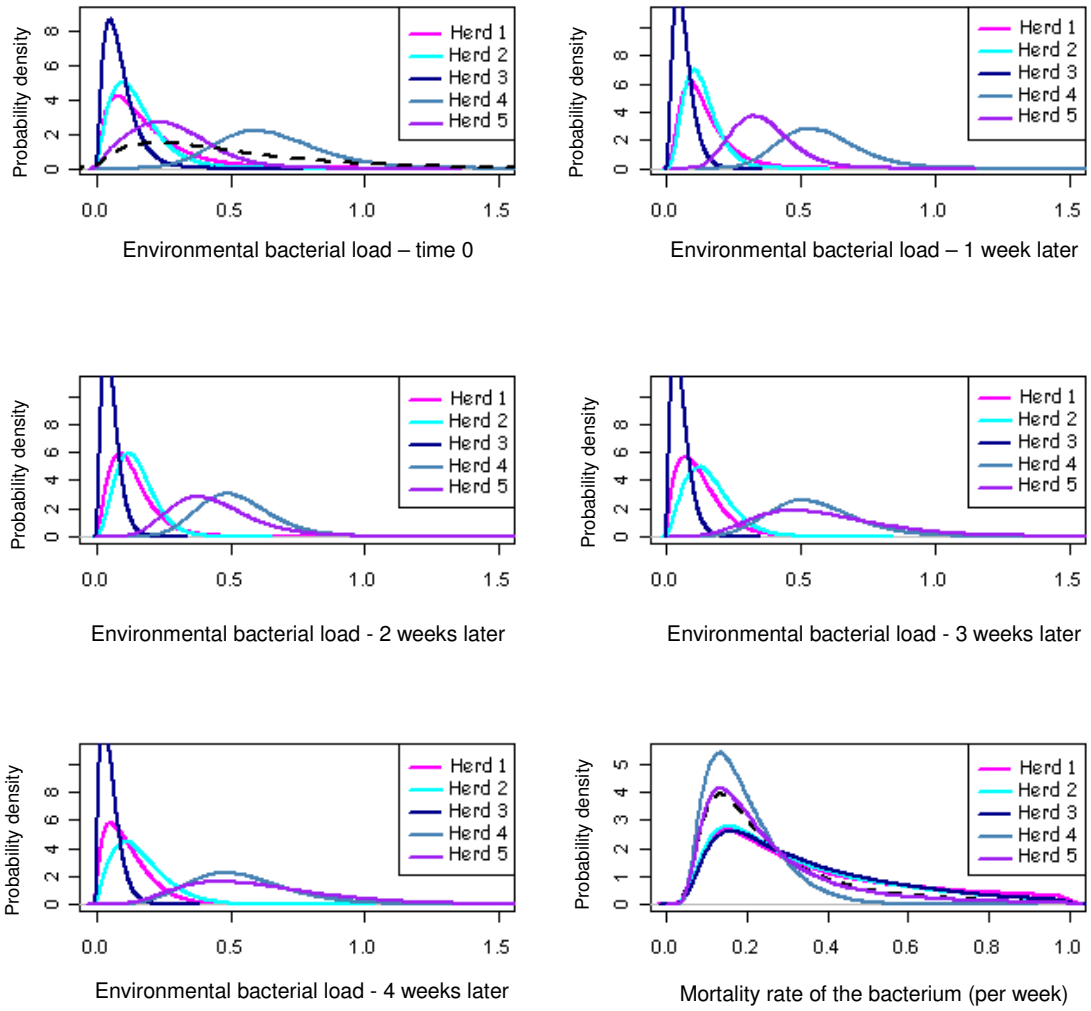

Figure 2.8. Posteriors of the environmental bacterial load (E) and of the mortality rate of the bacterium (μ). For the initial environmental bacterial load and for μ, priors are also drawn (black dotted line). For the initial environmental bacterial load, posterior distributions cover shorter intervals than those defined by prior distributions, which reveals that the data provide some information. This is not the case for μ.

f. Summary statistics for the initial real health states

At the prior level, the initial real health states, $R_{0,h}^{(i)}$, are random variables with a probability distribution specified by J, where $J_{x_j} = P(R_{0,h}^{(i)} = x_j)$ with $x_j \in \{x_1 = S, x_2 = I-, x_3 = I+, x_4 = R\}$.

Table 2.7 provides summary statistics of posterior distributions for marginal probability distributions of initial real health state.

Table 2.7. Priors and posteriors for the probability of initial real health states (J) in each of the five herds. For the marginal prior and posterior distributions, medians and 95% credible intervals (CI) are shown.

Model parameter	Prior distribution		Posterior median and 95% CI				
			Herd 1	Herd 2	Herd 3	Herd 4	Herd 5
Probability that the initial health state of cow *i* is *S*	Dirichlet (3.5,1,2.5,2)	0.380 (0.120 – 0.706)	0.327 (0.159 - 0.521)	0.439 (0.300 - 0.581)	0.688 (0.533 - 0.818)	0.524 (0.359 - 0.685)	0.306 (0.205 - 0.425)
Probability that the initial health state of cow *i* is *I-*		0.082 (0.003 – 0.369)	0.062 (0.004 - 0.194)	0.031 (0.001 - 0.109)	0.032 (0.001 - 0.134)	0.189 (0.071 - 0.347)	0.015 (0.000 - 0.064)
Probability that the initial health state of cow *i* is *I+*		0.261 (0.054 – 0.593)	0.237 (0.107 - 0.410)	0.140 (0.060 - 0.256)	0.079 (0.020 - 0.189)	0.143 (0.059 - 0.265)	0.336 (0.233 - 0.449)
Probability that the initial health state of cow *i* is *R*		0.203 (0.031 – 0.528)	0.351 (0.184 - 0.549)	0.377 (0.243 - 0.523)	0.183 (0.082 - 0.323)	0.129 (0.048 - 0.252)	0.333 (0.219 - 0.456)

CHAPTER 3

REPRESENTATION OF THE HETEROGENEITY OF SHEDDING IN THE MODEL OF WITHIN HERD SPREAD OF *C. BURNETII* AND IDENTIFICATION OF THE MOST INFLUENTIAL PARAMETERS OF THE INFECTION DYNAMICS

Picture : A. Senkowski

In the first section of this chapter, we will briefly describe how different population heterogeneities affect infection dynamics in many diseases, which are the implications for control purposes and how modelling accounts for these heterogeneities. In the second section, the variability of the shedding routes, duration and levels observed in data set A will be detailed. Then, some generalities will be provided on the approaches allowing identification of the most influential parameters of an infection dynamics. These key parameters are indeed of major importance: once identified, they have to be accurately assessed to improve both the model prediction and the understanding of processes involved in the infection spread; also, interventions impacting them are of great interest. The last section will describe our model of within herd spread of *C. burnetii* (with the representation of the individual variability of the shedding routes, duration and levels) and the sensitivity analysis performed. This part will be presented as it was submitted to Journal of Theoretical Biology.

I- Why and how to represent heterogeneity in host population?

When modelling the spread of an infection in a population, average quantities (e.g. average duration of infectiousness, average number of contacts with congenerics, average quantity of pathogen shed, etc.) are most of the time used as parameter values. This generic representation is acceptable as a first approach for providing a global view of the transmission process. However, populations are heterogeneous and individuals can have different physiological or behavioural characteristics, which are worthy to be taken into account. As an example, Diekmann & Heesterbeek [36] assume that in a population, the infectivity differs between individuals. Once epidemic growth takes off, all the values of infectivity are represented among the many infectious individuals and it is acceptable to work with the mean value of infectivity when describing the infection dynamics. However, it is not the case during the very first stages of infection where the values of infectivity of the few infected individuals have a great impact on the evolution of infection.

1. Two classic examples of heterogeneity in human diseases: sexually transmitted infections and childhood diseases

Host heterogeneity is well described in Sexually Transmitted Infections (STIs), for which high and low-risk individuals can be defined depending on their number of sexual contacts. By having many sexual partners, high-risk individuals have a higher risk of both contracting and transmitting the disease than low-risk ones. Therefore, models representing the spread of an STI should include several classes of individuals, and are more complex than models with assumed homogeneous host population [77]. One of the key parameters for direct-transmitted infections is the transmission rate, often denoted by β, defined as the rate per unit of time at which a susceptible, S and an infectious, I, individuals come into effective contact (i.e. a contact which leads to a new infection). This parameter can be seen as the product of the rate of contact between the two individuals and the probability that this contact will induce a new infection. In a standard *SIS* model where the population is assumed globally homogeneous (see chapter 2 section I for more details on different model structures), there is only one parameter β, whereas in an *SIS* model with two classes of individuals (high-risk and low-risk groups), there are four distinct β: transmission rate from high-risk individuals to high-risk ones, from high-risk to low-risk, from low-risk to high-risk, and from low-risk to low-risk. Thus, in an *SIS* model with two classes of individuals, there are more equations and parameters than in a standard *SIS* model. However, incorporating such heterogeneity in the model has several advantages: the infection prevalence can be determined for each of the different classes and used to define more efficient targeted control measures [163]. Besides, the basic reproduction ratio R_0[3] from structured models is generally larger than if the structures were ignored and all individuals had the same transmission rates [77].

Other well studied infections requiring partitioning of the host population are childhood diseases, such as measles or mumps. In this case, the distinction between classes is based on age rather than on the number of contacts with congenerics. Such diseases are common in childhood but rare for adults: indeed, in addition to an increased susceptibility, the

[3] The basic reproduction ratio (or basic reproductive number) is the expected number of secondary cases that a single infected individual will cause when introduced into a naïve population (i.e. a population with no immunity to the disease) and in the absence of control measures. When $R_0 < 1$, the infection will die out in the long run. When $R_0 > 1$, the infection will be able to spread in the population [36].

individuals who mix most with children (i.e. especially other children), are at the greatest risk [77]. In these models, the population is subdivided into a number of discrete compartments, classified depending on hosts' age and individuals progress through increasingly older age classes.

This type of models can also be used for animal diseases. As an example, Ferguson et al. [46] used an age-structured model when describing the spread of the prion responsible for the Bovine Spongiform Encephalopathy (BSE) through the cattle farms of the United Kingdom. Their model included many sources of heterogeneity: each cow was indexed by two variables, age and time-since-infection, on which transmission rates and susceptibility were dependent. Following the inclusion of this double dependence, the model became very complex. However, given the economic and public health importance of the BSE epidemic, it was crucial to achieve a high degree of accuracy [77].

2. Superspreading events occur in many infectious diseases

Large variations in infectiousness have been described for many infectious diseases, and especially for the Severe Acute Respiratory Syndrome (SARS). In the Singapore epidemic, of the first 201 probable cases reported, 103 were infected by just five source cases [89]. These individuals that directly infect a large number of other people are called superspreaders. The definition is here not age-related and the infectiousness[4] and susceptibility[5] of superspreaders seem not correlated, contrary to those of individuals infected by STIs. Lipsitch et al. [89] showed that the presence of superspreaders, and then the large variation in the effective reproduction number R[6] had a great influence on the early course of the epidemic: the variability in the effective reproduction number R means that many infected individuals transmit few or not at all while some transmit a lot. The probability that a single infected individual will result in a large epidemic is therefore

[4] The infectiousness of an individual describes its ability to transmit the infection to other hosts.

[5] The susceptibility of an individual describes its ability to get the infection from other hosts.

[6] The effective reproductive number R is the number of secondary cases generated by a single infected case once the epidemic is underway (i.e. the population is not fully susceptible). In the absence of control measures, $R = R_0 x$, where x is the proportion of the population susceptible. During the course of an epidemic, R declines because of the depletion of susceptibles in the population and the implementation of specific control measures. To stop an outbreak, R must be maintained below 1 [89].

lessened compared to the case where the value of *R* is the same on average but presents less variation. However, if the epidemic occurs, it can be very explosive.

The superspreading, although a key point in the 2003 SARS epidemic, was seldom represented in models until recently [48]. The 20/80 rule [169], which suggests that roughly 20% of the most infectious individuals are responsible for 80% of transmission, has been applied mainly to helminthic and sexually transmitting infections but not to other directly transmitted diseases. In 2005, Lloyd-Smith et al. [91] reassessed heterogeneous infectiousness. They considered that the infectiousness was distributed continuously in any population and that distinct homogeneous risk groups could not be defined a priori. In their model, the expected number of secondary cases caused by a particular infected individual (parameter equivalent to the *R* of Lipsitch et al. [89]) was drawn from a continuous probability distribution with population mean R_0, and superspreading events corresponded to realizations from the right-end tail of this distribution. Using contact tracing data from eight directly transmitted diseases, they showed high variation in individual infectiousness for most of the data sets. Model predictions accounting for this heterogeneity differed from average-based approaches, with disease extinctions more likely and outbreaks rarer but more explosive in the former case. Besides, control efforts targeting highly infectious individuals outperformed population-wide measures.

In a similar way, Matthews et al. [105] showed that British cattle infected by *Escherichia coli O157* was characterised by a high variability in bacterial shedding concentrations and consequently in infectiousness: a model assuming that all farms and all animals are governed by the same underlying dynamics was unable to explain the highly overdispersed distribution of prevalences of *Escherichia coli O157* shedding on Scottish farms [106]. The best fit to the prevalence data was obtained when incorporating variability in transmission rates at the animal level. This variability was both within host (i.e. variability over time for the same animal) and between hosts. In fact, 20% of the variance in bacterial counts could be attributable to host-to-host variation. Besides, the authors showed that 20% of the infections with the higher mean infectiousness contributed around 80% of the transmission. Effective control strategies would then consist of (i) targeting the super shedders (i.e. the most infectious individuals): preventing infection in 5% of the individuals with the highest mean infectiousness would bring R_0 below 1; and (ii) targeting bacterial carriage at high concentrations: limiting the bacterial load at 10^4 cfu/g (count corresponding to the top 6% of observed counts) would produce 48% of reduction in transmission, which would decrease R_0 below 1.

A different way to take into account superspreading was proposed by James et al. [69], through an event-oriented approach in which every individual had the potential of extensive spreading. Superspreading events (SSEs) were seen as stochastic consequences of environmental variability. James et al. [69] compared their model with the model of Lloyd-Smith et al. [91]: for most of the data sets, there was little difference in the Akaike information criterion[7], which illustrated that none of these models was clearly favoured over the other. The implications for control proposed by both groups of authors were different. For James et al. [69], as infections caused by non-SSEs could be relatively insignificant, targeted control policies based on reducing the frequency or severity of SSEs had to be implemented. The frequency of large gatherings of people or animals could be decreased by reducing the duration of working/school week or the frequency of animal markets. Moreover, to reduce the severity of SSEs, the maximum number of people (or animals) gathering together should be reduced. As these control measures did not require indentifying superspreaders, they were easier to implement than those proposed by Lloyd-Smith et al. [91] (i.e. targeting highly infectious individuals). However, the 'reality' of superspreading should lay somewhere between the event-oriented and individual-oriented approaches and modelling, both individual heterogeneity and rare SSEs being important challenges for the future [69]. It has to be highlighted that both models agreed about the consequences of superspreading phenomena: they cause less frequent but more explosive outbreaks. Garske et al. [50] also draw those conclusions when studying the impact of superspreading on patterns of disease outbreaks. Besides, these authors showed that outbreak sizes distributions were a less and less adapted guide to estimate R_0 of an infection as heterogeneity increases. Further studies on the extent and consequences of heterogeneity in infectiousness are then required.

We have just shown that identifying superspreaders would be useful [48]. However, such a task is very difficult to achieve in practice. As summarized by Lloyd-Smith et al. (Supplementary information of [91]), hosts, pathogens and environmental factors all contribute to variation of infectiousness. Contact rates are a key point: superspreaders are often noted to have high numbers of occupational or social contacts, or an activity that facilitates pathogen dispersion, such as food handling. Evolution of highly-transmissible pathogen strains is also possible although little studied. Besides, crowded or confined

[7] Akaike's information criterion is a measure of the goodness of fit of an estimated statistical model. The AIC is not a test of the model in the sense of hypothesis testing; rather it is a test between models - a tool for model selection. Given a data set, several competing models may be ranked according to their AIC, with the one having the lowest AIC being the best.

settings, as well as the delay before an infectious patient is isolated, have a strong influence on individual infectiousness. Lastly, host-pathogen interactions affect transmission rates via variation in symptom severity and in pathogen load or shedding. Identifying factors such as age, genetic, diet or other management factors that might lead to high levels of shedding would then be of great interest [105].

II- The heterogeneity of shedding in *C. burnetii* infections

In cattle herds infected by *C. burnetii*, shedding routes are often not concomitant and the titres in *C. burnetii* are highly variable between shedders. In addition, some cows, mostly highly-seropositive, shed in milk with a persistent shedding pattern [59]. Based on this knowledge, our model was rendered more realistic by representing this special type of shedders (called $I+^{milk\ pers}$) and the shedding routes and levels for each shedder type (*I-*, *I+* and $I+^{milk\ pers}$). The partitioning of the population into these different categories was made on a probabilistic basis and the values of the discrete probability distributions controlling it were based on observations from data set A presented in Chapter 1.

This section provides some details on the observed distributions of shedding routes and levels for *I-*, *I+* and $I+^{milk\ pers}$ cows in order to (i) highlight, if any, differences between those three types of shedders and (ii) feed the mathematical model. Besides, as the uterus and mammary glands of females are sites of chronic *C. burnetii* infection [107], a second objective was to determine if the calving had an impact on these distributions. Therefore, we separately analysed the data regarding the cows which calved in the month before the sampling and the data of those which calved more than a month before.

1. Shedding routes

According to our data set, seven shedding route categories were defined: shedding in (i) milk only ("Milk"), (ii) vaginal mucus only ("Muc"), (iii) faeces only ("F"), (iv) milk and mucus ("Milk+Muc"), (v) milk and faeces ("Milk+F"), (vi) mucus and faeces ("Muc+F"), (vii) milk, mucus, and faeces ("all routes").

The variability in the shedding routes was noticed both within cow (i.e. over time for a given cow) and between cows. Over 47 *I-* and *I+* cows observed shedders twice one week

apart, only 53.2% were allocated at the 2nd time of shedding to the same shedding route category as the 1st time. Almost half of them were shedders in milk only; the other half were shedders is mucus only. Only one cow was observed shedding in milk and mucus twice one week apart.

As shown in Figure 3.1, there is a significant difference between the shedding route distributions of *I-* and *I+* individuals, for cows which calved more than a month before (Fisher test, p-value < 0.001): *I-* cows mostly shed in mucus only (43% of cases) and milk only (34% of cases) whereas *I+* animals shed preferentially in milk only (61% of cases). For cows which calved in the month before the sampling, there is no significant difference between *I-* and *I+* cows.

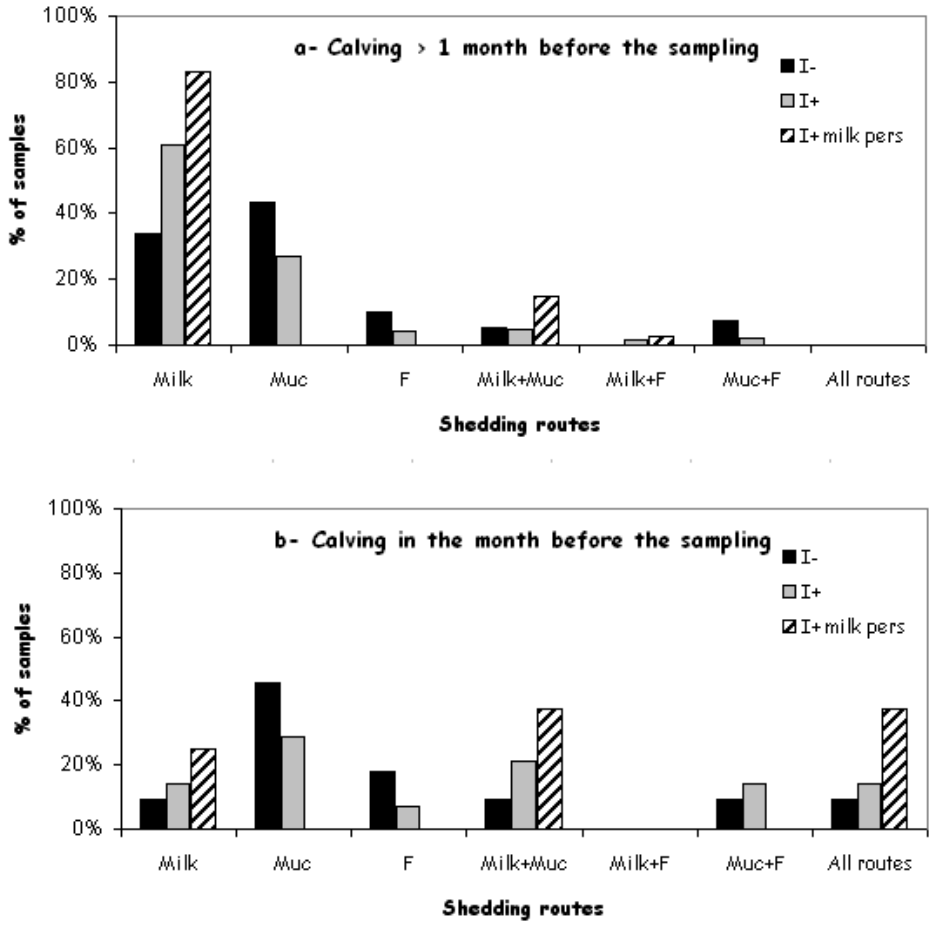

Figure 3.1. Distribution of the shedding routes with respect to the type of *I* cow.
In black: *I-* (11 samples for the recently calved cows, 97 for the other ones); in grey: *I+* (14 samples for the recently calved cows, 151 for the other ones);in hatched: *I+ milk pers* (8 samples for the recently calved cows, 82 for the other ones)

For both *I+* and *I+ milk pers* individuals, there is a significant difference between the shedding route distributions of cows which calved more than a month before and cows

which calved in the month before the sampling (Fisher tests, p-value < 0.001): cows which recently calved, shed less often in milk only. There is no significant difference between these two types of cows for *I-* individuals.

Thus, based on the analysis of our data, we chose for the *I+* and *I+ milk pers* animals different probability distributions for cows which calved more than a month before the sampling and for cows which calved in the month before the sampling. Therefore, we used five different probability distributions in the model for the shedding routes (see Table 3.1 of section IV): one for the *I-* cows, two for the *I+* cows, and two for the *I+ milk pers* cows.

2. Shedding levels

In the real-time PCR, the quantification is relative and based on the Ct (cycle threshold) of an endogenous internal positive control, the GAPDH. Since for the faeces samples, there are not enough cells, an exogenous positive control is used and no quantification is performed. Therefore, only shedding levels in milk samples and vaginal swabs are presented.

Like for the shedding routes, the observed variability in the shedder levels is both within and between individuals. However, the former is less frequent: over 33 cows shedding through the same route twice one week apart, 69.7% shed the 2nd time in the same shedding level category as the 1st time.

The distributions of the shedding levels for the different types of *I* are presented Figure 3.2. Most of the *I-* individuals shed at low titres, whatever the shedding route and moment of calving. For the *I+* individuals which calved more than a month before, the shedding level distribution in milk samples significantly differs (i) from the one in mucus samples for the same kind of cows, and (ii) from the one in milk samples for cows which calved in the previous month (Fisher tests, p-value < 0.001). There is no significant difference between the shedding level distributions in milk and mucus samples for recently calved cows, whatever the type of *I*. Most of *I+ milk pers* individuals shed in mid titres except recently calved cows shedding in mucus (Fisher test, p-value < 0.001). These latter more often shed in low titres.

According to the descriptive statistics analysis of our dataset, five different probability distributions for the shedding levels were considered in the model (see Table 4.1 in section IV): as the probability distributions for all the *I-* and for the *I+* mucus shedders which calved more than a month before did not differ significantly, the same probability

distribution *Q1* was used for these types of animals. Different probability distributions *Q2* and *Q3* were respectively defined for the *I+* milk shedders which calved more than a month before, and for the *I+* cows which calved in the previous month whatever their shedding route. Lastly, we used a probability distribution *Q4* for the $I+^{milk\ pers}$ mucus shedders which calved more than a month ago and a probability distribution *Q5* for all the $I+^{milk\ pers}$ milk shedders which calved more than a month ago and for the $I+^{milk\ pers}$ which calved in the previous month.

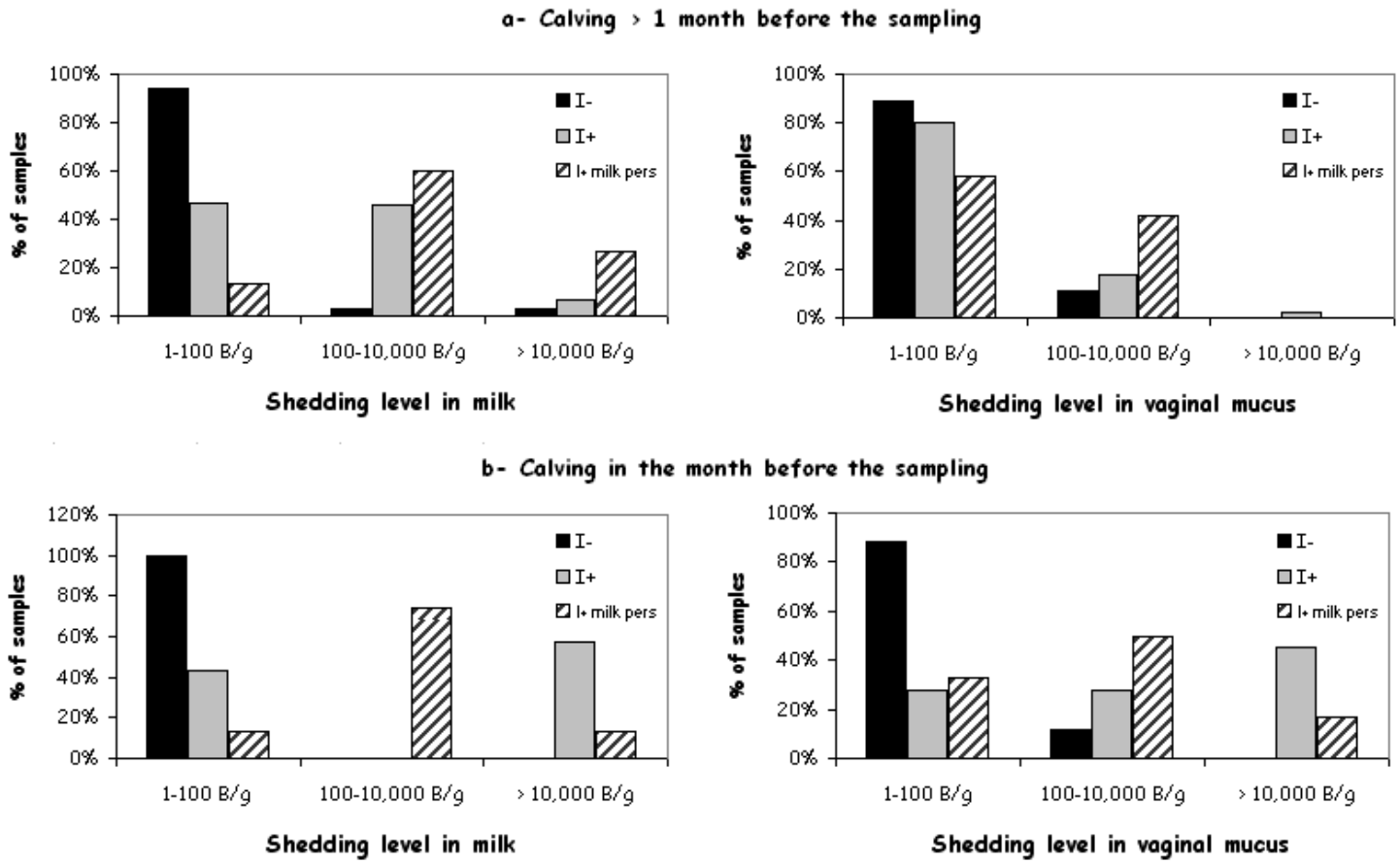

Figure 3.2. Distribution of the shedding levels function of different infectious statuses of: (a) cows which calved more than a month before the sampling, (b) cows which calved in the month before the sampling. In black: $I-$ (3 samples in milk and 8 samples in mucus for the recently calved cows, 38 samples in milk and 47 in mucus for the other cows), in grey: $I+$ (7 samples in milk and 11 samples in mucus for the recently calved cows, 101 samples in milk and 51 in mucus for the other cows), and in hatched black: $I^{+\,milk\,pers}$ (8 samples in milk and 6 samples in mucus for the recently calved cows, 82 samples in milk and 12 in mucus for the other cows)

III- Why and how to perform a sensitivity analysis?

When building a model, the modeller has often several possibilities of model structures to answer his questions. It is then crucial to understand the impact of specific parameterizations on the outputs. Concerning model parameters, two situations may schematically occur: either many different values or no quantitative information are available in the literature or from expert opinions. Although ideally this uncertainty could be reduced by collecting more data, this is not always possible in practice. It seems then essential to investigate the way uncertainties of different orders propagate on the outputs variables [30], especially if the model aims at advising policy makers.

These two aspects can be explored through sensitivity and uncertainty analysis. Sensitivity analysis "is the study of how the variation in the output model can be apportioned, qualitatively or quantitatively, to different sources of variation, and of how the given model depends upon the information fed into it" [140]. It allows ordering by importance the strength and relevance of the inputs when studying the variation in the output. The uncertainty analysis quantifies the uncertainty in the outcome of a model. In other words, sensitivity analysis determines the relationships between information flowing in and out of a model. And in this sense it should be distinguished from the uncertainty analysis which quantifies the variability of the output due to the incomplete knowledge of the system but does not link this variability to the variability of the different inputs.

1. Aims of sensitivity analyses

Sensitivity analysis has a wide range of goals [140]. First, it allows determining if the model has the expected behaviour. If the model is strongly dependant on *a priori* non influential factors[8] or, conversely, if the variation of *a priori* highly influential factors has no impact on the model outputs, there is a need to revise the model structure or parameter values. Sensitivity analysis also allows to define the most important factors, which, if fixed to their most likely value, would lead to the greatest reduction in the variance of the output [30]. This is called factors prioritization setting and helps prioritize research needs in terms, for instance, of data acquisition. Besides, sensitivity analysis can be useful to simplify the model: the non influential factors can be either eliminated from

[8] "Factor" is defined as any input included in the sensitivity analysis. It can be a parameter, an input variable, or a module of the model [140].

the model or fixed to any value of their domains without significantly increasing the output variability. This latter point is called factors fixing setting. At last, sensitivity analysis allows determining the region of the space of inputs factors with the largest model variation and to detect interactions between factors.

2. How to perform sensitivity analysis?

The first step is to determine which input factors will be considered for the analysis. This choice depends on the question in study and on the available knowledge on the modelled system. Then, for each input factor, the range of variation and either the factor levels (i.e. the possible factor values within the variation range) or the factor probability distribution should be defined. The third step consists of generating factor combinations which will be used as inputs when running the model. This step is a crucial one and many methods (not detailed here) are available to design experiments. For example, if the factors are defined through probability distributions, the selection of samples from these distributions can be made randomly or by Latin Hypercube sampling[9]. When factors are defined by discrete levels, complete or fractional factorial designs can be used[10]. The fourth step is to run the model and then to determine the value of the output of interest for each combination of input factors. At last, the influence or relative importance of each input factor on the output variable has to be assessed. Different methods are available and the choice is not easy as each technique has its strengths and weaknesses.

3. Types of methods[11]

When the model is computationally expensive or has a large number of input factors, screening methods are useful. They allow identifying the factors that control most of the output variability but are only qualitative: the input factors are ranked by order of importance but the methods do not quantify the relative difference between factors.

[9] The range of each input factor is divided into *N* intervals of equal marginal probability *1/N*, and for each input factor one point is generated in each interval. There are then *N* non-overlapping realizations for each of the input factors [140].

[10] A full factorial experiment is an experiment whose design consists of two or more factors, each with discrete possible values or "levels", and whose experimental units take on all possible combinations of these levels across all such factors. If the number of combinations in a full factorial design is too high to be logistically feasible, a fractional factorial design may be done. In this case some of the possible combinations are omitted, according to specific rules established to render feasible the estimation of desired effects (main effect, second order interactions, etc.)

[11] The description of the following methods is based on Cariboni et al. [30] and Saltelli et al. [140].

Typical screening designs are *one-at-a-time* (OAT) experiments: the effect of the variation of a single factor is estimated keeping all the others fixed at their estimated values. This type of method does not allow estimating factor interactions as only one factor varies at each time. An exception is the OAT design proposed by Morris: the experiment covers the entire space over which the factors may vary, whereas in standard OAT experiments, the factors vary only around their nominal values (Figure 3.3). For each factor i at the given point x of the sampling space is calculated an elementary effect defined as $d_i(x) = \frac{y(x_i,\ldots,x_{i-1},x_i+\Delta,x_{i+1},\ldots,x_k)-y(x)}{\Delta}$ with y the output and $x=(x_1, x_2, \ldots, x_k)$ a selected point in the sampling space. The mean of the elementary effects of a given factor measures its overall effect on the output while the standard deviation accounts for interactions.

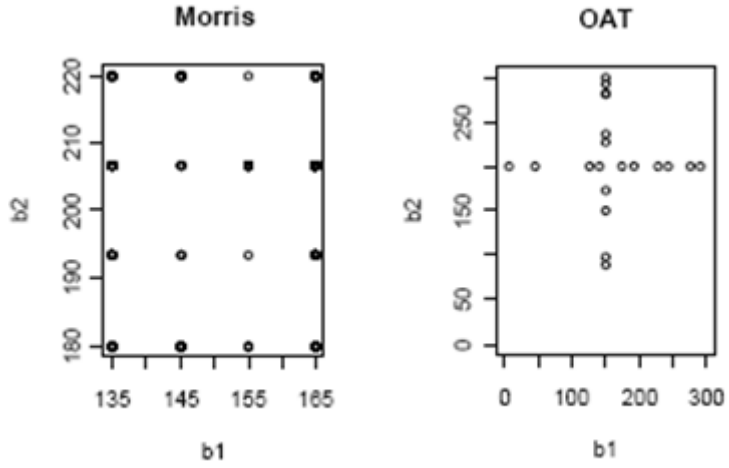

Figure 3.3. Space sampling in grid for Morris OAT and in cross for standard OAT; b1 and b2 are two inputs factors.

Local analysis is a quantitative method which usually consists of calculating partial derivatives of the output functions with respect to the input factors. This method takes into account only small variations around the factor nominal values and is usable only for linear models.

Global approaches allow assessing the effect of an input factor on the output variation when all the other input factors are varying. They are suitable for non linear and/or non additive models and allow measuring the sensitivity over the entire range of each input parameter. The output variance can be decomposed in order to impute to each input factor its contribution. Different techniques exist: a factorial decomposition of the model

variance by an analysis of variance (ANOVA), Sobol's decomposition, and Fourier Amplitude Sensitivity Test (FAST). The two latter are widely-used even computationally costly. They allow quantifying V_i , the amount of output variance explained by each input factor *i*. Sensitivity indices S_i, representing the main effect of factor *i*, can then be defined as the ratio $V_i / V(Y)$, with *V(Y)* the output variance. When interactions are also considered, a total effect index ST_i can be calculated to account for all the contributions to the output variation due to factor *i* (its main effect plus all its interactions). The ANOVA also allows calculating sensitivity indices but with this method, the sensitivity analysis is based on an approximation of the model by a simpler linear model [113]. As an example with two input factors Z_1 and Z_2, the response variability can be decomposed as follows:

$$\underbrace{\sum_{a\,b}\left(\hat{Y}_{ab}-\mu\right)^2}_{SS_T} = \underbrace{m\sum_{a}\alpha_a^2}_{SS_1} + \underbrace{m\sum_{b}\beta_b^2}_{SS_2} + \underbrace{\sum_{a,b}\gamma_{a,b}^2}_{SS_{12}} .$$

SS_T: total variability in the model responses

SS_1 and SS_2: sum of squares associated respectively with the main effect of Z_1 and Z_2

SS_{12} : sum of squares associated with the interaction between Z_1 and Z_2

$\hat{Y}_{ab}$ denotes the model response when z_1 = a and z_2 = b, $\mu = \hat{Y}_{\bullet\bullet}$ is the general mean, *m* is the number of possible values for a and b, $\alpha_a = \hat{Y}_{a\bullet} - \mu$ is called the main effect of factor Z_1 when z_1 = a et $\beta_b = \hat{Y}_{\bullet b} - \mu$ is the main effect of factor Z_2 when z_2 = b and $\gamma_{a,b}$ is the interaction related effect. The main effects sensitivity indices are therefore defined as $S_1 = \frac{SS_1}{SS_T}$ and $S_2 = \frac{SS_2}{SS_T}$, the interaction sensitivity indices are $S_{12} = \frac{SS_{12}}{SS_T}$ and the total sensitivity indices are $TS_1 = \frac{SS_1 + SS_{12}}{SS_T}$ and $TS_2 = \frac{SS_2 + SS_{12}}{SS_T}$.

Although sensitivity analyses are widely-used for deterministic models, adaptations of these methods for stochastic models are still in progress and no consensus on the steps to be followed is currently available. One of the possible approaches is to consider as outputs the mean and standard deviation of the output of interest over the repetitions of the model. Besides, all the previously described methods are well-defined for non dynamical outputs. However, in epidemiological models, it seems useful to identify the factors most influencing the entire dynamics of infection. Global sensitivity analysis could be applied separately on each time point of each output, but successive dates enclose relative

redundant information and also interesting features of the dynamic may be missed out. In order to jointly consider all the points of time series in the sensitivity analysis, Lamboni et al. [82, 84] developed a new method based on principal component analysis and on analysis of variance. A generalized sensitivity index is computed for each model parameter. The proposed index synthesizes the influence of the parameter on the whole time series output. As described in the following section, this is the approach we adopted for the sensitivity analysis of the model of heterogeneity of shedding in *C. burnetii* spread in a herd that we developed.

IV- Manuscript: Modelling the effect of heterogeneity of shedding on the within herd *Coxiella burnetii* spread and identification of key parameters by sensitivity analysis

Aurélie Courcoul[1,2,*], Hervé Monod[3], Mirjam Nielen[4], Don Klinkenberg[4], Lenny Hogerwerf[4], François Beaudeau[1,2,5], Elisabeta Vergu[3]

[1]INRA, UMR1300 Bio-agression, Epidémiologie et Analyse de Risque, Nantes, France
[2]Oniris, UMR1300 Bio-agression, Epidémiologie et Analyse de Risque, Nantes, France
[3]INRA, UR341 Mathématiques et Informatique Appliquées, Jouy-en-Josas, France
[4]Utrecht University, Faculty of Veterinary Medicine, Utrecht, The Netherlands
[5]Université Nantes, Angers, Le Mans, France

Submitted to Journal of Theoretical Biology

1. Abstract

Coxiella burnetii is the bacterium responsible for Q fever, a worldwide zoonosis. Ruminants, especially cattle, are recognized as the most important source of human infections. Although a great heterogeneity between shedder cows has been described, no previous studies have determined which features such as shedding route and duration or the quantity of bacteria shed have the strongest impact on the environmental contamination and thus on the zoonotic risk. Our objective was to build a model representing the spread of *C. burnetii* within a dairy cattle herd, taking into account the heterogeneity of shedding and to identify key parameters whose variation highly influences the infection dynamics.

We proposed an individual-based stochastic model in discrete time describing the evolution of the infection representing both the individual variability of the shedding duration, routes and intensity as well as herd demography. To compare the influence of the epidemiological parameters on different temporal outputs, we performed a sensitivity analysis consisting of a Principal Component Analysis followed by an ANOVA. Our findings showed that the most influential parameters were the probability distribution governing

the levels of shedding, especially in vaginal mucus or faeces, the characteristics of the bacterium in the environment (i.e. its survival and the fraction of bacteria shed reaching the environment), and some physiological parameters related to the intermittency of shedding (transition probability from a non shedding infected state to a shedding state) or to the transition from one type of shedder to another one (transition probability from a seronegative shedding state to a seropositive shedding state).

Our study seemed crucial for the understanding of the infection dynamics. As control measures should impact the parameters influencing the infection dynamics most, our model can now be used to assess the effectiveness of different control strategies for *C. burnetii* infection within dairy cattle herds.

2. Introduction

Q fever is a worldwide zoonosis caused by *Coxiella burnetii*. This intracellular bacterium infects a wide range of animals and is associated with reproductive disorders in domestic ruminants [5, 20, 26, 98]. Goats, sheep and cattle are recognized as the main source of human infection [96, 109, 142, 164]. Infected animals shed bacteria through various routes (parturition products, faeces, urine, vaginal mucus, milk) [10, 20, 57]. As the bacterium survives very well in the environment, humans can get infected by inhaling contaminated dusts or aerosols. This was recently experienced in the Netherlands where more than 3,000 cases were reported since 2007 [159]. Although Q fever is asymptomatic in humans in more than 60% of cases, it can lead to acute or chronic infections and cause flu-like syndrome, hepatitis, pneumonia, endocarditis or abortions [52, 128]. Hence, for public health and economic and animal health concerns, it is important to control *C. burnetii* infections in livestock herds.

In *C. burnetii* infections, a great heterogeneity between shedders has been described [15, 37, 131]: the shedding duration and routes, as well as the level of shedding (i.e. the quantities of bacteria shed) are variable between cows. According to Guatteo et al. [59], cows can shed sporadically or persistently, the shedding routes are rarely concomitant and the concentrations of bacteria shed in vaginal mucus or milk can vary from less than 100 Bacteria/g to more than 1,000,000 B/g. Heterogeneity of shedding is known to affect infection dynamics in many diseases [104] but it is generally difficult to determine which of its aspects are the most influential. The length of shedding, its route, the quantity of bacteria shed, or other features may all have the strongest impact on the environment contamination by *C. burnetii* and thus on the zoonotic risk in the case of Q fever infection.

A representation of the disease spread within a herd as well as the identification of key parameters characterizing the heterogeneity of shedding are thus critical for the understanding of the infection dynamics. In addition, the effectiveness of a control measure was shown to be dramatically improved by targeting the individuals transmitting the pathogen most (e.g. in the case of *Escherichia coli* O157 infection [105], of measles epidemics [50] or of *Salmonella* transmission [86]). However, understanding and predicting the spread of *C. burnetii* in a herd or identifying such key parameters cannot be assessed by field experiments alone. In this context, mathematical models are useful tools for understanding how the infection spreads within the herd and how various inputs (such as epidemiological characteristics of infected animals) affect the dynamics [97]. Techniques such as sensitivity analysis allow assessing the impact of the uncertainty and variability in the parameters on models outputs and hence determining key factors [140]. It consists in studying how the variation in the outputs of the model can be apportioned to different sources of variation, and how the model depends upon the information fed into it.

The aim of our study is first to build a model representing the spread of *C. burnetii* within a dairy cattle herd, taking into account the heterogeneity of shedding and second to determine the key parameters related to this heterogeneity whose variation highly influences the infection dynamics. The model that we will present is, to our knowledge, the first one proposed in the literature for Q fever spread coupling epidemiological aspects (mainly heterogeneity in shedding) with herd demography. The sensitivity analysis that will be described, followed by the presentation and the discussion of the results, is an original approach allowing dealing with temporal outputs.

3. Model

a. General description

The epidemic model that we developed describes the spread of *C. burnetii* within a dairy cattle herd, considering different health statuses, which are defined by excretion of bacteria, immunity and various characteristics related to the shedding route and the quantity of bacteria delivered in the environment (Figure. 3.4 and Table 3.1 for the parameter description). The herd demography is included through interaction of lactation and gestation statuses with shedding. The model is stochastic, individual-based and in discrete time with a time step of one week, which is appropriate for both epidemiological and herd management processes. The stochasticity has two main sources: for each individual, all the transitions between health states are supposed stochastic and the

quantities of bacteria shed in the environment follow discrete distributions, different according to the shedding route.

- ✓ **Health states and their associated transitions**

Each cow is in one of the six mutually exclusive health states at a given time (Figure. 3.4): *S* (susceptible, non-shedder without antibodies), *I-* (shedder without antibodies), *I+* (shedder with antibodies), $I^{+\,milk\ pers}$ (shedder with antibodies, shedding in milk at higher levels and for a longer period of time than *I+*, as described by Guatteo et al. [59]), *C+* (non-shedder with antibodies), *C-* (non-shedder without antibodies which was infected and had antibodies in the past). All shedding cows *I* are subdivided according to their shedding routes: (1) I_1, milk only, (2) I_2, vaginal mucus and/or faeces, (3) I_3, both.

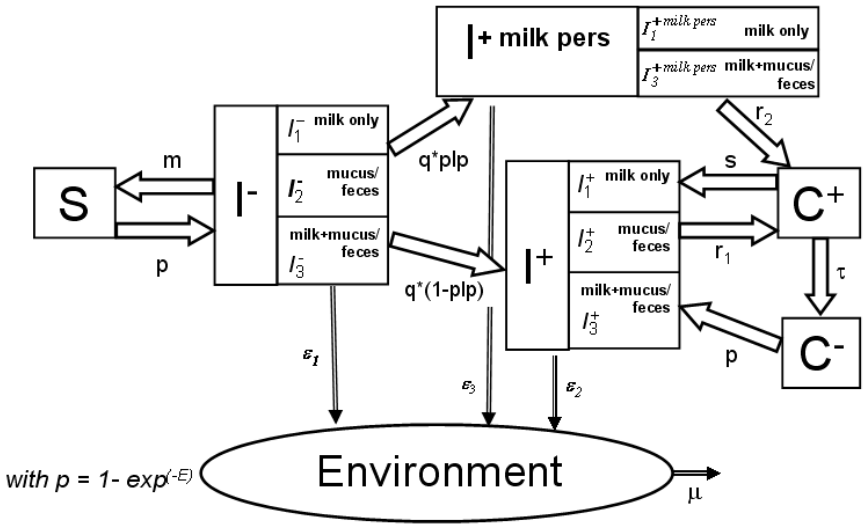

Figure 3.4. Flow diagram describing the modelled spread of *C. burnetii* within a cattle herd. The health states are: *S*, susceptible, non-shedder cow without antibodies, *I-*, shedder cow without any antibodies, *I+*, shedder cow with antibodies, $I+^{milk\ pers}$ shedder cow with antibodies shedding in milk in a persistent way, *C+*, non-shedder cow with antibodies and *C-*, non-shedder cow without antibodies which was infected and had antibodies in the past. *I-* and *I+* cows are in the shedding route category *1* if they shed in milk only, *2* if they shed in vaginal mucus/faeces only and *3* if they shed in milk and vaginal mucus/faeces. $I+^{milk\ pers}$ cows are in the shedding category *1* if they shed in milk only and 3 if they shed in milk and vaginal mucus/faeces. *E* represents the environmental bacterial load. The model parameters are presented in Table 3.1. ε_1, ε_2 and ε_3 are the quantities of bacteria shed during a time step by an individual *I-*, *I+* and $I+^{milk\ pers}$ respectively and contaminating the environment. These quantities are the sum of all quantities of bacteria shed by all the shedders through all the shedding routes *Qty*, times ρ the fraction of bacteria shed reaching the environment of the herd.

Table 3.1. Parameters of the epidemiological model: description, standard values and values tested in the first experiment of the sensitivity analysis

Factor name	Description	Standard value	Values tested in the sensitivity analysis			
m (week^{-1})	Transition probability $I-$ => S	0.7[a]	0.33	0.8		
q (week^{-1})	Transition probability $I-$ => $I+$	0.02[a]	0.01	0.2		
pIp	Proportion of cows going from $I-$ to $I+$ and becoming $I+^{milk\ pers}$	0.5	0.2	0.7		
r_1 (week^{-1})	Transition probability $I+$ => $C+$	0.2[a]	0.11	0.33		
r_2 (week^{-1})	Transition probability $I+^{milk\ pers}$ => $C+$	0.02	0.01	0.06		
s (week^{-1})	Transition probability $C+$ => $I+$	0.15[a]	0.04	0.4		
τ (week^{-1})	Transition probability $C+$ => $C-$	0.0096				
μ (week^{-1})	Mortality rate of *C. burnetii**	0.2[a]	0.08	0.5		
probav	Probability of abortion after a transition S => $I-$, $C+$ => $I+$ or $C-$ => $I+$	0.02				
ρ^{mf}	Proportion of bacteria shed through mucus/faeces filling the environment compartment	0.2	0.05	0.15	0.3	0.5
ratio ρ^{milk}/ρ^{mf}	Where ρ^{milk} = proportion of bacteria shed through milk filling the environment compartment	0.125	0.0625	0.125	0.25	0.5

α	$\alpha_{1,}$ milk	Probability distribution of the shedding routes for the $I-$ cows	0.31[b]	0.31	0.6	0.06	0
	$\alpha_{2,,}$ mucus/faeces		0.62[b]	0.62	0.3	0.75	1
	$\alpha_{3,}$ milk+mucus/faeces		0.07[b]	0.07	0.1	0.19	0
β	$\beta_{1,}$ milk	Probability distribution of the shedding routes for the $I+$ cows after 4 weeks post-calving	0.61[b]	0.61	0.14	0.2	0.77
	$\beta_{2,}$ mucus/faeces		0.33[b]	0.33	0.5	0.7	0.09
	$\beta_{3,}$ milk+mucus/faeces		0.06[b]	0.06	0.36	0.1	0.14
β_{calv}	β_{calv1}, milk	Probability distribution of the shedding routes for the $I+$ cows in the 4 first weeks post-calving	0.14[b]	0.14	0.61	0.2	0
	β_{calv2}, mucus/faeces		0.5[b]	0.5	0.33	0.7	0.33
	β_{calv3}, milk+mucus/faeces		0.36[b]	0.36	0.06	0.1	0.67
γ	γ_{1}, milk	Probability distribution of the shedding routes for the $I+^{milk\ pers}$ cows after 4 weeks post-calving	0.83[b]	0.83	0.25	0.5	1
	γ_{3}, milk+mucus/faeces		0.17[b]	0.17	0.75	0.5	0
γ_{calv}	γ_{calv1}, milk	Probability distribution of the shedding routes for the $I+^{milk\ pers}$ cows in the 4 first weeks post-calving	0.25[b]	0.25	0.83	0.5	0
	γ_{calv3}, milk+mucus/faeces		0.75[b]	0.75	0.17	0.5	1
Q1	low level	Probability distribution of the shedding levels for all the $I-$ and for the $I+$ shedding in mucus/faeces after 4 weeks post-calving	0.85[b]	0.85	0.6	0.25	0.15
	mid level		0.15[b]	0.15	0.4	0.25	0.6
	high level		0[b]	0	0	0.5	0.25
Q2	low level	Probability distribution of the shedding levels for the $I+$ shedding in milk after 4 weeks post-calving	0.4[b]	0.4	0.85	0.25	0.15
	mid level		0.5[b]	0.5	0.15	0.25	0.6
	high level		0.1[b]	0.1	0	0.5	0.25

Q3	low level	Probability distribution of the shedding levels for all the *I+* in the 4 first weeks post-calving	0.25[b]	0.25	0.85	0.6	0.15
	mid level		0.25[b]	0.25	0.15	0.4	0.6
	high level		0.5[b]	0.5	0	0	0.25
Q4	low level	Probability distribution of the shedding levels for the $I+^{milk\ pers}$ shedding in mucus/faeces after 4 weeks post-calving	0.6[b]	0.6	0.85	0.25	0.15
	mid level		0.4[b]	0.4	0.15	0.25	0.6
	high level		0[b]	0	0	0.5	0.25
Q5	low level	Probability distribution of the shedding levels for all the $I+^{milk\ pers}$ shedding in milk and for the $I+^{milk\ pers}$ in the 4 first weeks post-calving	0.15[b]	0.15	0.85	0.6	0.25
	mid level		0.6[b]	0.6	0.15	0.4	0.25
	high level		0.25[b]	0.25	0	0	0.5

[a]from Courcoul et al. [35]

[b]calibrated to match field data (R. Guatteo 2009, personal communication)

[*]includes both the natural mortality of the bacterium and its removal in relation to the periodic cleaning of the cattle housing

As shown in Figure 3.4 and matrix $\mathbf{M_t}$ below, by inhaling bacteria contained in the environment, a susceptible cow, *S*, can become infectious, *I-*, with probability *p* (expressed at each time step as $1 - \exp^{(-E_t)}$ where E_t is the quantity of bacteria in the environment of the herd at time *t*). Either it manages to eliminate the bacterium and becomes (apparently) *S* again (transition probability *m*) or it produces antibodies and continues shedding (transition probability *q*). It can then become *I+* or *I+ milk pers* with rate *pIp*. When it stops shedding, it becomes *C+* (transition probabilities r_1 and r_2 respectively). Since the shedding can be intermittent as observed in experimental and field studies [59, 131], a transition from *C+* to *I+* is assumed (transition probability *s*). A *C+* individual can also clear the infection, loose its antibodies and become *C-* (transition probability *t*). If this individual is infected again, its humoral immunity is assumed to be immediately reactivated and it becomes *I+* again (without passing through the *I-* state) with the same probability as an individual in state *S*.

- ✓ **Heterogeneity of shedding**

Both shedding routes and levels of shedding are taken into account in our model. We assume that the probability distribution corresponding to the assignment to one of the three categories of shedding routes defined above is different for each infectious state *I-*, *I+* or *I+ milk pers* (probability distributions denoted by α, β and γ). As the quantification of *C. burnetii* is not available in the faeces samples, the distribution of the associated titers of bacteria is assumed to be similar to the distribution of the titers in the mucus samples. Besides, we assume that shedding in vaginal mucus or in faeces have the same impact in terms of contamination of the environment (same ρ equal to ρ^{mf}, ρ being the fraction of bacteria shed reaching the environment of the herd). Therefore, these two excretion routes are gathered into a single category. Concerning the shedding levels, three categories are represented: low, moderate and high level shedding, corresponding respectively to a quantity of bacteria *Qty* of 1/3000, 1/30 and 1 unit of environment. The probability to shed at one of these levels (represented by the probability distributions *Q*, described in Table 3.1) depends on the infectious state (*I-*, *I+* or *I+ milk pers*) and on the shedding route (milk or mucus/faeces). Both the distributions *Q* and the ratios between the *Qty* were determined based on field data (R. Guatteo 2009, pers. comm.).

The quantity of bacteria arriving into the environment during a time step represents the sum of *Qty* times ρ for all the shedders releasing bacteria through all the shedding routes. This last parameter is assumed to be lower for shedding in milk, ρ^{milk}, than for shedding in

mucus/faeces, ρ^{mf} (i.e. a lower proportion of the bacteria shed in milk is supposed to arrive into the environment of the herd, because most of the milk is directly sent to the bulk, and then to the dairy industry). It has to be stressed here that animals in the third category (I_3^-, I_3^+ and $I_3^{+\,milk\,pers}$ respectively) are assumed to contribute through two simultaneous shedding routes to the filling up of the environment, namely milk and mucus/feces.

- ✓ **Infection dynamics**

The temporal dynamics of the individual health states is modelled using Markovian transitions. Let $R_t^{(i)} \in \{S,\ I_1^-,\ I_2^-,\ I_3^-,\ I_1^+,\ I_2^+,\ I_3^+,\ I_1^{+\,milk\,pers},\ I_3^{+\,milk\,pers},\ C+,\ C-\}$ be the health state of individual i at time t. $R_t^{(i)}$ depends on $R_{t-1}^{(i)}$ and on E_t, the quantity of bacteria in the environment at time t. The transition probabilities can be contained in the matrix $\mathbf{M}_t=(m_{t,jk})$:

$$M_t=\begin{pmatrix}
1-p_t & p_t\alpha_1 & p_t\alpha_2 & p_t\alpha_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
m & (1-m-q)\alpha_1 & (1-m-q)\alpha_2 & (1-m-q)\alpha_3 & q(1-pIp)\beta_1 & q(1-pIp)\beta_2 & q(1-pIp)\beta_3 & q\,pIp\,\gamma_1 & q\,pIp\,\gamma_3 & 0 & 0\\
m & (1-m-q)\alpha_1 & (1-m-q)\alpha_2 & (1-m-q)\alpha_3 & q(1-pIp)\beta_1 & q(1-pIp)\beta_2 & q(1-pIp)\beta_3 & q\,pIp\,\gamma_1 & q\,pIp\,\gamma_3 & 0 & 0\\
m & (1-m-q)\alpha_1 & (1-m-q)\alpha_2 & (1-m-q)\alpha_3 & q(1-pIp)\beta_1 & q(1-pIp)\beta_2 & q(1-pIp)\beta_3 & q\,pIp\,\gamma_1 & q\,pIp\,\gamma_3 & 0 & 0\\
0 & 0 & 0 & 0 & (1-r_1)\beta_1 & (1-r_1)\beta_2 & (1-r_1)\beta_3 & 0 & 0 & r_1 & 0\\
0 & 0 & 0 & 0 & (1-r_1)\beta_1 & (1-r_1)\beta_2 & (1-r_1)\beta_3 & 0 & 0 & r_1 & 0\\
0 & 0 & 0 & 0 & (1-r_1)\beta_1 & (1-r_1)\beta_2 & (1-r_1)\beta_3 & 0 & 0 & r_1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-r_2)\gamma_1 & (1-r_2)\gamma_3 & r_2 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-r_2)\gamma_1 & (1-r_2)\gamma_3 & r_2 & 0\\
0 & 0 & 0 & 0 & s\beta_1 & s\beta_2 & s\beta_3 & 0 & 0 & 1-s & \tau\\
0 & 0 & 0 & 0 & p_t\beta_1 & p_t\beta_2 & p_t\beta_3 & 0 & 0 & 0 & 1-p_t
\end{pmatrix}$$

where $m_{t,jk}=P(R_t^{(i)}=x_k \mid R_{t-1}^{(i)}=x_j)$ for $t=1,\ldots,Tmax$, $k,j=1..11$, $\{x_1=S,\ x_2=I_1^-,\ x_3=I_2^-,\ x_4=I_3^-,\ x_5=I_1^+,\ x_6=I_2^+,\ x_7=I_3^+,\ x_8=I_1^{+\,milk\,pers},\ x_9=I_2^{+\,milk\,pers},\ x_{10}=C+,\ x_{11}=C-\}$. The probabilities of categorical distributions α, β and γ verify the conditions $\alpha_1+\alpha_2+\alpha_3 = 1$, $\beta_1+\beta_2+\beta_3 = 1$ and $\gamma_1+\gamma_3 = 1$.

As the environmental bacterial load at time t, E_t, is dependent on E and the prevalences of shedders ($I_{1,t}^-, I_{2,t}^-, I_{3,t}^-, I_{1,t}^+, I_{2,t}^+, I_{3,t}^+, I_{1,t}^{+\,milk\,pers}, I_{3,t}^{+\,milk\,pers}$) at time $t-1$, the environment dynamics can be expressed by the equation: $E_{t+1}=(1-\mu)E_t+\sum_{k,l}\left(\rho^k Qty_l \sum_{i,j} n_{t,ijkl}\right)$, where $i \in \{I-,\ I+,\ I+^{milk\,pers}\}$, $j \in$ {≤ 4 weeks post calving, > 4 weeks post calving}, $k \in$ {milk, mucus/faces}, $l \in$ {low, medium, high},

$\rho^{milk} = \rho^{mf} * ratio$ and $n_{t,ijkl} \sim Multin(N_{t,ijk}, Q_{c(i,j,k)})$. N_{ijk} represent numbers of animals in corresponding health states at time t and $Q_{c(i,j,k)}$ are the probability distributions governing shedding levels, which are not necessary distinct for each state (for instance, as explained above, based on field data, we assumed that $Q_{c(I^-,>4 \text{ weeks post calving, mucus / faeces})} = Q_{c(I^+,>4 \text{ weeks post calving, mucus / faeces})} \equiv Q_I$).

- ✓ **Herd demography**

The epidemic model is coupled with a model of herd demography. Only cows (neither heifer nor calf) are represented in our model. No lactating cow is purchased by the farmer. Thus, only S primiparous cows which have just calved (former heifers becoming lactating cows) are assumed to enter the herd. These introductions of animals can occur at any time of the year. However, if at time t, the size of the herd is above 1.15 times the initial size, we assume that no heifer is introduced at this time.

The culling rate depends on the lactation number. The culling of animals can occur at any time of the year. However, if at time t, the size of the herd is below 0.85 times the initial size, it is assumed that no cow is culled at this time.

For each cow, we represent the lactation/gestation cycle. We consider a calving-calving interval of 55 weeks. The lactation cycle is composed of 47 weeks of lactation starting at calving followed by 8 weeks of dry period. The gestation cycle is composed of a non gestation period of 15 weeks starting at calving followed by a gestation of 40 weeks.

Table 3.2. Description of the parameters of the herd demography model and their standard values.

Description		Standard value
Replacement rate (year^{-1})		0.355
Culling rate (week^{-1})	lactation 1	0.0057
	lactation 2	0.0052
	lactation 3	0.0065
	lactation 4	0.0067
	lactations 5&6	0.0161
Probability distribution at time 0 for the lactation numbers of the cows	lactation 1	0.337
	lactation 2	0.252
	lactation 3	0.173
	lactation 4	0.11
	lactation 5	0.088
	lactation 6	0.04

- ✓ **Interactions between epidemiological processes and herd demography**

Dry cows can not become $I+^{milk\ pers}$ and a $I+^{milk\ pers}$ cow becoming dry can stay either in $I_1^{+\ milk\ pers}$ or in $I_3^{+\ milk\ pers}$ but she is assumed not to shed any bacteria into the environment through milk. A dry cow becoming or staying $I-$ or $I+$ is necessarily in sub category 2 (shedding in vaginal mucus and/or faeces). As shown in Table 3.1, the date of calving also impacts the probability distributions (α, β, γ and Q) of shedding sub categories for $I+$ and $I+^{milk\ pers}$ cows.

Regarding the abortions, cows in gestation can abort during the 3 weeks following infection or resumption of shedding (which can occur during a transition from S to $I-$, from $C+$ to $I+$ or from $C-$ to $I+$). Abortions can occur at any time of the gestation. It is assumed that a cow can abort only once in her life. If a cow aborts in the first or second third of gestation, she sheds at that moment a moderate quantity of bacteria in the mucus/faeces shedding route, whereas if the abortion occurs in the last third of gestation (late abortion), she sheds a high quantity of bacteria in the same shedding route. In addition, if a cow aborts in the first or second third of gestation, the non gestation period is reduced to 8 weeks (instead of 15 weeks after a normal calving or a late abortion). If a cow aborts after the week 22 of gestation, it starts a new lactation. If she aborts before, her current lactation continues for a maximum of 50 weeks of lactation. Afterwards, she is dried off.

At last, from mid-March to mid-November, we assume that cows in lactation and dry cows are not kept all together. Therefore, two types of environment are defined: $E_{building}$ is the environmental bacterial load of the main buildings and close pastures for lactating cows, while E_{dry} is the environmental bacterial load of the specific pastures of dry cows. The probability of infection is thus different for lactating and dry cows during this period ($p_{building}$ and p_{dry} respectively). Outside this period, there is only one compartment environment in which all the cows (lactating and dry) shed their bacteria. This environment is $E_{building}$.

b. Initial conditions and parameter values of the standard scenario

At t=0, the herd consists of 50 cows. To initiate the infection cycle, a primiparous $I+$ cow which has just calved is introduced into a wholly susceptible herd.

The epidemiological parameters are put at their standard values (Table 3.1): parameters m, q, r_1, s and μ come from a study where they were estimated through Bayesian inference using data from five French chronically infected dairy cattle herds [35]; probability distributions of shedding related parameters, α, β, β_{calv}, γ, γ_{calv}, *Q1, Q2, Q3, Q4* and *Q5,* were qualitatively calibrated to match field data (R. Guatteo 2009, personal communication). The parameters governing the demography and herd management (Table 3.2) were chosen to represent a standard French dairy cattle herd.

To account for the variability in Q fever infections, 200 repetitions of the same scenario were run over a 5-year simulation period.

4. Sensitivity analysis

a. Outputs and factors

We conducted a sensitivity analysis to identify the parameters that mostly contributed to the output variability. Various scenarios were run, each of them being characterized by a specific combination of parameter values, in order to relate the variability obtained for the outputs to that induced by the input parameters. Eight outputs were considered (Table 3.3): *(i)* $E_{building}$, *(ii)* E_{dry}, *(iii)* the prevalence of milk shedders, *(iv)* the prevalence of mucus/faeces shedders, *(v)* the prevalence of shedders in milk in a persistent way, *(vi)* the seroprevalence, *(vii)* the number of abortions per herd per year, and (*viii)* the extinction rate. All these outputs except the number of abortions and the extinction rate were computed weekly over a 5-year period.

Parameters related to the herd demography were fixed at their nominal values of Table 3.2 since demography and herd management processes are considered as well known. The sensitivity of the model outputs was evaluated with respect to the epidemiological parameters, which are those given in Table 3.1, except for τ. These 19 parameters are thus the inputs of the sensitivity analysis and they will be called factors in the rest of the paper. They belong to two categories: parameters concerning the transitions between health states and the environment (m, q, pIp, r_1, r_2, s, μ) and parameters directly related to the heterogeneity in shedding (i.e. α, β, β_{calv}, γ, γ_{calv}, *Q1, Q2, Q3, Q4, Q5,* ρ^{mf} and *ratio* $\left(\rho^{milk}/\rho^{mf}\right)$). The parameters in the first category were estimated from field data previously [24], but some uncertainty still remains (due for instance to the limitation of the data). They were included in the sensitivity analysis but with relatively limited ranges

of variation. In this study we focused on the latter category of parameters because they directly describe heterogeneity related aspects, which represented our main objective.

Table 3.3. Description of the outputs of the sensitivity analysis.

Output name	Description
$E_{building}$	Environmental bacterial load of the main buildings and close pastures
E_{dry}	Environmental bacterial load of the specific pastures for the dry cows
Prevalence of milk shedders	Proportion of animals of the herd shedding in milk, i.e. $\frac{I_1^- + I_3^- + I_1^+ + I_3^+ + I_1^{+\,milk\ pers} + I_3^{+\,milk\ pers}}{N}$ (*)
Prevalence of mucus/faeces shedders	Proportion of animals of the herd shedding in vaginal mucus and/or faeces, i.e. $\frac{I_2^- + I_3^- + I_2^+ + I_3^+ + I_3^{+milk\ pers}}{N}$ (*)
Prevalence of shedders in milk in a persistent way	Proportion of animals $I^{+\,milk\ pers}$, i.e. $\frac{I_1^{+\,milk\ pers} + I_3^{+milk\ pers}}{N}$ (*)
Seroprevalence	Proportion of animals with antibodies, i.e. $\frac{I_1^+ + I_2^+ + I_3^+ + C^+ + I_1^{+\,milk\ pers} + I_3^{+\,milk\ pers}}{N}$ (*)
Number of abortions per herd per year	
Extinction rate	Proportion of runs of a particular scenario leading to an extinction of the infection (**)

* N denotes the herd size

** the infection is assumed extinct when there is no I and C left until the end of the simulation time

b. Design of experiments

All the designs were generated using R 2.10.1 [127] and PLANOR R package [27].

- ✓ **First experiment**

We used a fractional factorial experiment design, with four parameter values (called levels) per factor related to the shedding and two levels for the other parameters (values in Table 3.1). As our model is stochastic, we ran the model for each combination of factor levels 30 times. Since the complete factorial design would lead to too many combinations (exactly $30 \times 4^{12} \times 2^7$ simulations), a fractional factorial design of resolution V was chosen. Such a design allows estimating the main effects and two-factor interactions, provided higher order interactions are assumed to be negligible [18, 80]. In the present case, a design was obtained with 4,096 scenarios. Thus, we ran 122,880 realizations of the model (i.e. 30 repetitions for each of the 4,096 scenarios).

- ✓ **Second experiment**

A complete factorial design for the eight most influential factors according to the first experiment was performed. This enabled us to more accurately quantify the impact of the interactions between these eight factors and also to disentangle potential confounded main effects and interactions. Besides, we determined in this experiment the factors that mostly contributed to the variability of the extinction rate between repetitions. The remaining 11 parameters were put to their standard value (Table 3.1). For this second study, we ran 2,048 scenarios with 30 repetitions each.

- ✓ **Third experiment**

In a third analysis, the influence of factors *Q* was specifically explored. Since the probability distributions *Q* depend on the type and route of shedding, this analysis enabled us to explore which type of shedders (*I-*, *I+* or $I+^{milk\ pers}$) and which type of shedding route (milk or mucus/faeces) played a major role in the variability of the outputs. Thus, the probability distributions of the shedding levels were varied independently, which generated 10 factors. Probability distributions *Q** were recorded as follows: for milk and mucus/faces respectively, *Q1** and *Q2** refer to the distributions of the shedding levels for the *I-*, *Q3** and *Q4** refer to those for the *I+* after 4 weeks post-calving, *Q5** and *Q6** are similar to the two former but correspond to the 4 first weeks post-calving, *Q7** and *Q8** refer to the distributions of the shedding levels for the $I+^{milk\ pers}$ after 4 weeks post-calving, and finally *Q9** and *Q10** are the symmetric of *Q7** and *Q8** for the 4 first weeks post-calving. Thus, former factor *Q1* of first and second experiments corresponds to new factors *Q1**, *Q2** and *Q4**, former factor *Q2* to *Q3**, former factor *Q3* to *Q5** and *Q6**, former factor *Q4* to *Q8** and former factor *Q5* to *Q7**, *Q9** and *Q10**.

A fractional factorial design for the 10 new factors *Q** with four levels each was generated. These four levels were (0.85, 0.15, 0), (0.6, 0.4, 0), (0.25, 0.25, 0.5) and (0.15, 0.6, 0.25) for the probability to be in (low, mid, high) shedding level respectively. The other parameters were put to their standard values given in Table 3.1. For this third experiment, we ran 1,024 scenarios with 30 repetitions each.

c. Analysis of the temporal outputs (of the first, second and third experiments)

In order to compare the influence of factors on the seven outputs which exhibit temporal dynamics (all outputs except the extinction rate), we applied a method developed by

Lamboni et al. [82] and used by Lurette et al. [95] to identify key parameters influencing *Salmonella* infection dynamics in a pig batch. The results are recorded as tables with one row for each scenario and one column for each output time points (260 weekly time points for the first six outputs and five annual time points for the abortion number).

This method allows simultaneously analyzing potentially correlated variables (here the successive time points of a given output). It consists in two main steps. First, a Principal Component Analysis (PCA) is operated in order to provide linear combinations (or components) of the initial variables (here the columns of our tables) explaining the maximum of inertia (i.e. variability) between scenarios. Only the first three principal components (PC) were kept since they are sufficient to cover most variability amongst simulations. The PCA provides to each line of the tables a score on each component. The second step involves an ANOVA, including the main effects and the two-factor interactions for all factors and carried out on the scores of each of the components considered. Sensitivity indices (SI), corresponding to the main effect or to interactions, and total sensitivities (TS), corresponding to the sum of the main effect and the interactions, were calculated for each factor and for each component. This analysis was performed with R 2.10.1 [127] and multisensi R package [83].

The analyses were performed on both the mean and standard deviation of the 30 repetitions of each scenario, in order to assess the two sources of variability influencing the outputs: the model intrinsic stochasticity and the parameter variability generated by the factorial designs.

d. Analysis of the extinction rate (of the second experiment)

An ANOVA was performed to assess the influence of the eight most influential factors on the extinction rate. It was calculated for each scenario defined by the complete factorial design of the second experiment.

e. Analysis of the outputs at a the time point 260 (of the first, second and third experiments)

In order to determine the factors with the highest influence on the output variability as a whole, we performed a joint analysis on the values of the six dynamic outputs (first six lines of Table 3.3) at the last simulation time step (week 260). This time point was chosen to illustrate the long-term steady-state of the system. Thus, the two-step analysis (PCA followed by ANOVA) was performed twice on six output variables. The first analysis was

done on the mean and the second on the standard deviation of the 6 dynamic outputs at time 260.

5. Results

a. Infection dynamics of the standard scenario

Over the 200 repetitions of the standard scenario, 37 led to the extinction of infection (defined as the absence of animals in I or C states in the herd) occurring on average in week 56 after the introduction of the initial infected cow (min: week 11, max: week 171). The mean seroprevalence and the mean prevalences of shedders increased with time (Figure 3.5) to reach respectively 34.7% on average [0 – 57.1% for the percentiles 2.5% and 97.5% respectively] and 35.5% [0 – 61.7% for the percentiles 2.5% and 97.5% respectively] five years after the initial infection. The ratio between the mean prevalence of shedders and the mean prevalence of milk shedders was around 2.5 in the first weeks of simulation, then it decreased to reach 1.84 at the end of the simulation time. The mean environmental bacterial load $E_{building}$ increased with time corresponding to a mean transition probability from S to I-, $p_{building}$, equal to 0.43 at the end of the simulation time. On the contrary, the mean environmental bacterial load E_{dry} was close to 0 for the 5 years of simulation (results not shown). The median abortion number was equal to 2 per herd per year the first year and 3 per herd per year afterwards, but a large variability surrounded these values [0-9 for the percentiles 2.5% and 97.5% respectively]. In addition, as shown in Figure 3.6, the route and the level of shedding of a shedder cow had a great impact on the contamination of the environment. This result is an unsurprising consequence of the model parameterization. As expected, the most common low level shedding category did not contribute much to the increase of the environmental bacterial load. On the contrary, shedders in mucus/faeces of the mid level category and shedders in milk of the high level category filled the environment in a non negligible way. Above all, shedders in mucus/faeces of the high level category (both non aborting and aborting cows) had the greatest impact.

b. Influence of the epidemiological factors on the model outputs

The results obtained with 30 runs for each parameter set were robust: the mean and the percentiles 2.5, 50 and 97.5 of our outputs were similar to those obtained with 200 runs (results not shown).

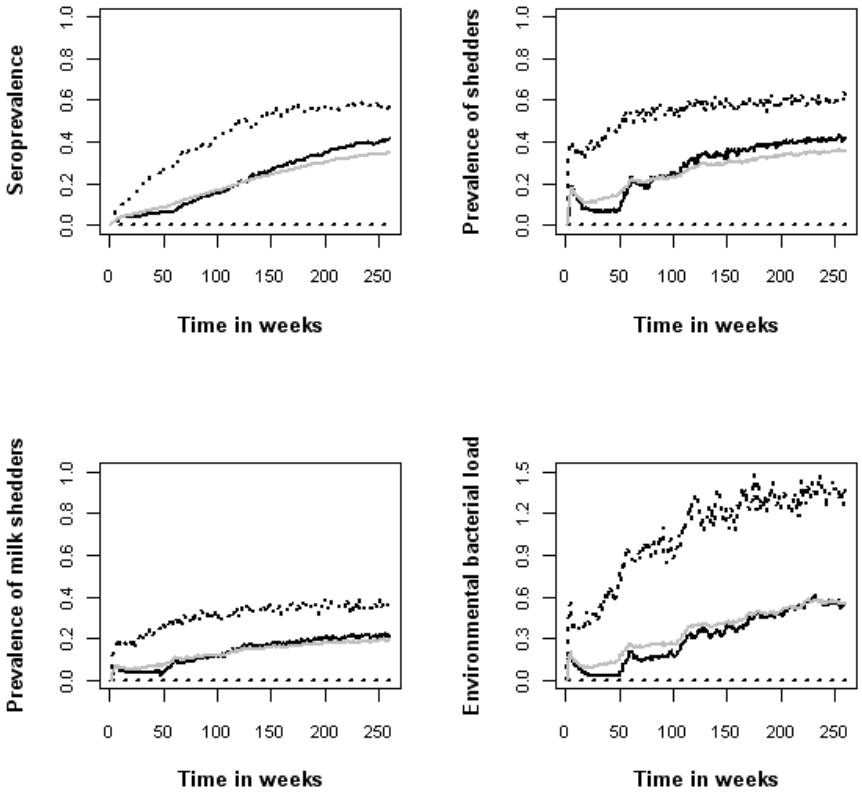

Figure 3.5. Temporal dynamics of the seroprevalence, prevalence of shedders, prevalence of milk shedders and environmental bacterial load $E_{building}$: mean (grey plain line), median (black plain line) and percentiles 2.5 and 97.5% (black dotted lines).

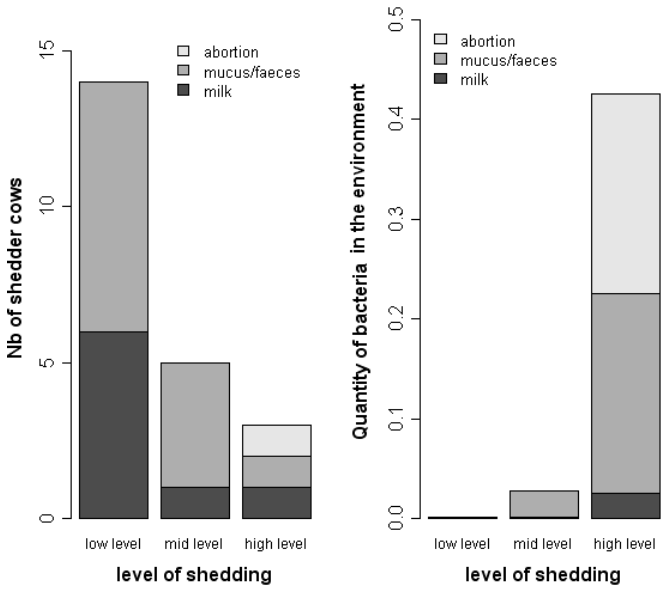

Figure 3.6. Number of shedders for each shedding route and each shedding level and their contributions in terms of contamination of the environment. Example according to the results from a given run at a given time.

- ✓ **First experiment**

As shown in Table 3.4, since the inertia obtained for the first PC was very high for each of the model outputs (except in the joined analysis), only the results on the first PC are presented. For the two mean environmental bacterial loads, the factors *Q1* (the probability distribution of the shedding levels for all the *I-* and for the *I+* shedding in mucus/faeces after 4 weeks post-calving), μ (the mortality rate of *C. burnetii*) and ρ^{mf} (the proportion of bacteria shed through mucus/faeces filling the environment compartment) were the most influential ones. For the mean prevalences of mucus/faeces and milk shedders, the most sensitive factors were *q* (the transition probability from *I-* to *I+*), *s* (the transition probability from *C+* to *I+* representing the intermittency of shedding) and *Q1*, whereas the mean prevalence of milk shedders in a persistent way was mostly impacted by *pIp* (the proportion of cows going from *I-* to *I+* and becoming $I+^{milk\ pers}$), *q* and r_2 (the transition probability from $I+^{milk\ pers}$ to *C+*). Concerning the mean seroprevalence, the factor *q* had a TS higher than 60%. Lastly, the most influential factors of the mean abortion number were *q*, *Q1*, *s*, μ and ρ^{mf}. Globally, the most sensitive two-factor interactions (with a SI higher than 5%) were *Q1:q* on the variability of the abortion rate, *Q1:*, ρ^{mf}*:Q1*, ρ^{mf}*:*μ on the variability of the environmental bacterial loads and *q:pIp, q:*r_2*, pIp:*r_2 on the variability of the prevalence of milk shedders in a persistent way.

Concerning the variability of the standard deviations of the outputs, the same factors as above were identified as the most influential ones for the environmental bacterial loads, the prevalence of milk shedders in a persistent way and the abortion number. The main effect of the factors was always very low (no SI higher than 5%) on the prevalences of mucus/faeces shedders, whereas the part of two-factor interactions was much more important. The most sensitive factors were *Q1*, *q*, ρ^{mf} and Q_3 (the probability distribution of the shedding levels for all the *I+* in the 4 first weeks post-calving). Regarding the standard deviation of the prevalence of milk shedders, the most sensitive factors were *q*, *Q1* and α (the probability distribution of the shedding routes for the *I-* cows).

For the joined analysis on six of the dynamic outputs at time 260, the inertia obtained for the first PC was much lower and the second PC had to be taken into account. For the means, the most influential factors were *Q1*, *q*, *s*, μ and ρ^{mf} on the first PC and *Q1*, μ, ρ^{mf} and *q* on the second one, by order of importance. For the standard deviation, the most sensitive factors were *q*, *Q1*, r_2 and Q_3 on the first PC and *Q1*, ρ^{mf} and μ on the second.

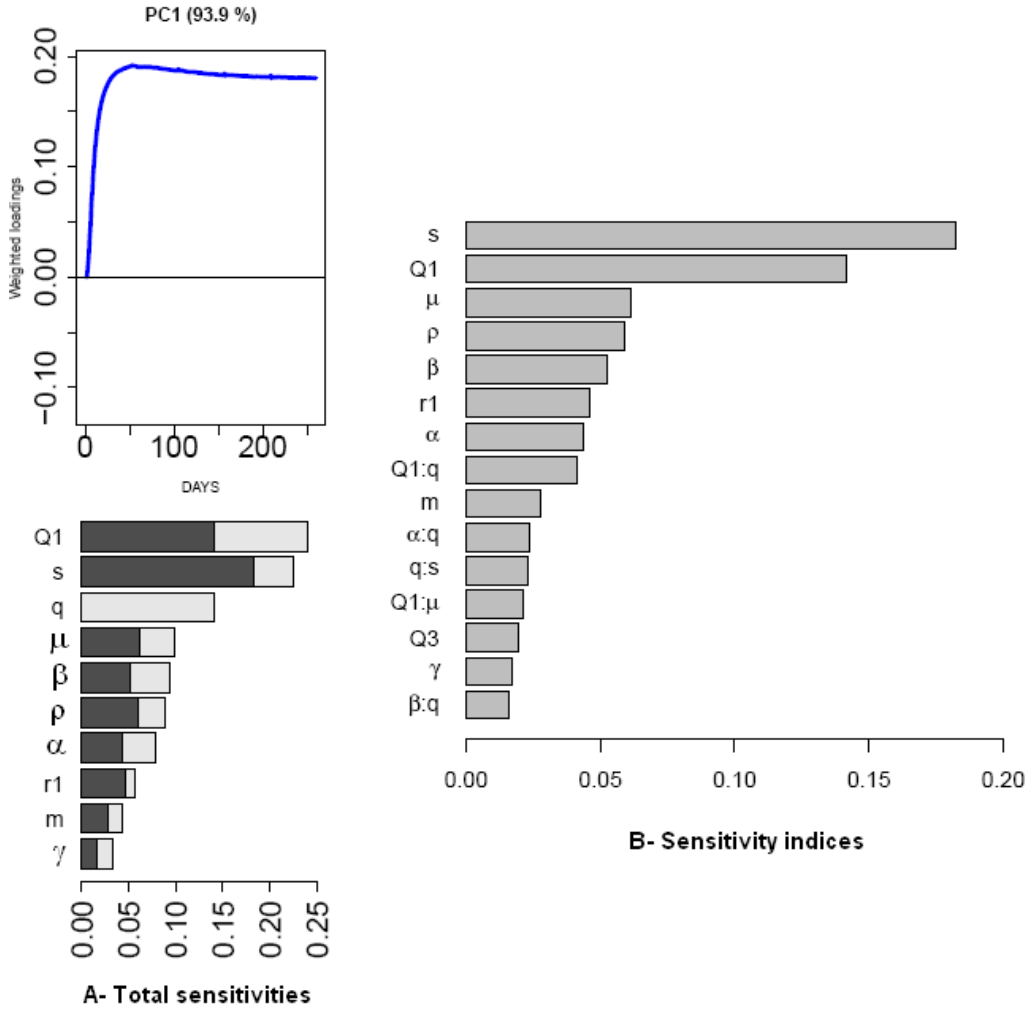

Figure 3.7. Sensitivity analysis on the mean prevalence in mucus/faeces shedders over time: results of the ANOVA performed for the first component (inertia: 93.9%). (A) Loadings defining the principal component for each time variable (in abscissa) and total sensitivities for the 10 most influential factors ranked in descending order. Sensitivities are split in main effect (black) and two-factor interactions (grey). (B) Sensitivity indices of the 15 main factorial terms (main effects or interactions) in descending order.

Table 3.4. Results of the first experiment of the sensitivity analysis: for the mean and standard deviation of each output, the proportion of total inertia represented by the first principal component and the corresponding most sensitive factors are given. The factors whose total sensitivities are higher than 10% for the component received one cross, those whose total sensitivities are greater than 25% received two crosses and those whose total sensitivities are greater than 50% received three crosses.

	Environment $E_{building}$		Environment E_{dry}		Prevalence of milk shedders		Prevalence of m/f shedders		Prevalence of milk shedders in a persistent way		Seroprevalence		Abortion rate		Joined analysis time point 260	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
inertia	99.1%	94.2%	98.1%	95.1%	94.3%	84.8%	93.9%	85.0%	93.1%	87.7%	97.1%	86.4%	93.5%	87.3%	48.7%	43.6%
m																
q					++	+	+	+	++	++	+++	++	++	++	+	+
pIp									++	++						+
r_1																
r_2									++	++						+
s					+		+					+	+	+	+	
μ	++	++	++	++									+	+	+	
ρ^{mf}	++	++	++	++				+					+	+	+	
$ratio$ ρ^{milk}/ρ^{mf}																
α						+										
β																
β_{calv}																
γ																
γ_{calv}																
$Q1$	++	++	++	++	+	+	+	+			+		++	++	+	+
$Q2$																
$Q3$								+				+				+
$Q4$																
$Q5$																

- ✓ **Second experiment**

The eight factors chosen in the second experiment (appearing as the most influential according to the findings of the first experiment) were *Q1*, *Q3*, *q*, *s*, r_2, μ, ρ^{mf} and *pIp*. The results obtained were globally similar to the results of the first experiment described above (same most influential factors, sometimes in a slightly different order), suggesting that no important interactions were confounded with the main effects in the first analysis.

Besides, the most sensitive factors on the extinction rate were firstly *Q1* and μ (with a SI higher than 14%) then *q*, *s*, and *Q3* (with a lesser SI, but higher than 5%). The most sensitive two-factor interactions were *Q1:q* and *Q1:* μ.

- ✓ **Third experiment**

Amongst the probability distributions of the shedding levels, *Q2**, characterizing the mucus/faeces seronegative shedders was globally the most influential factor on the variability of the outputs (Table 3.5 and Figure 3.8). *Q4** and *Q6**, the probability distributions of the shedding levels for mucus/faeces seropositive shedders at anytime also had a significant impact and, to a lesser extent, *Q1**, the probability distribution of the shedding levels associated to the milk seronegative shedders and *Q8**, the probability distribution of the shedding levels for persistent milk shedders excreting in mucus/faeces. Moreover, the only interaction among the five most sensitive terms was *Q2*:Q4**. Overall, the factors with the greatest impact were probability distributions of the shedding levels in mucus/faeces.

6. Discussion

In this study, we proposed the first model of *C. burnetii* spread within a dairy cattle herd taking into account the individual variability of shedding, defined in duration, routes and intensity. Simulated infection dynamics are consistent with field data: at the last time point of the simulated time series (five years after the introduction of the initial infectious case), the mean seroprevalence is around 35% [23.3% – 47.8% for the 25^{th} and 75^{th} percentiles respectively], which is consistent with the mean observed seroprevalence in cows (mean: 40%, 25^{th} and 75^{th} percentiles: 25% and 51% respectively) of 56 naturally infected French dairy herds [149]. At the same simulated time point, the mean prevalence of shedders is around 35.5% [0 – 61.7% for the percentiles 2.5^{th} and 97.5^{th} respectively] whereas in the field, the apparent proportion of shedder cows is 45.5% in Guatteo et al. [57] and 38.9% in Rodolakis et al. [131].

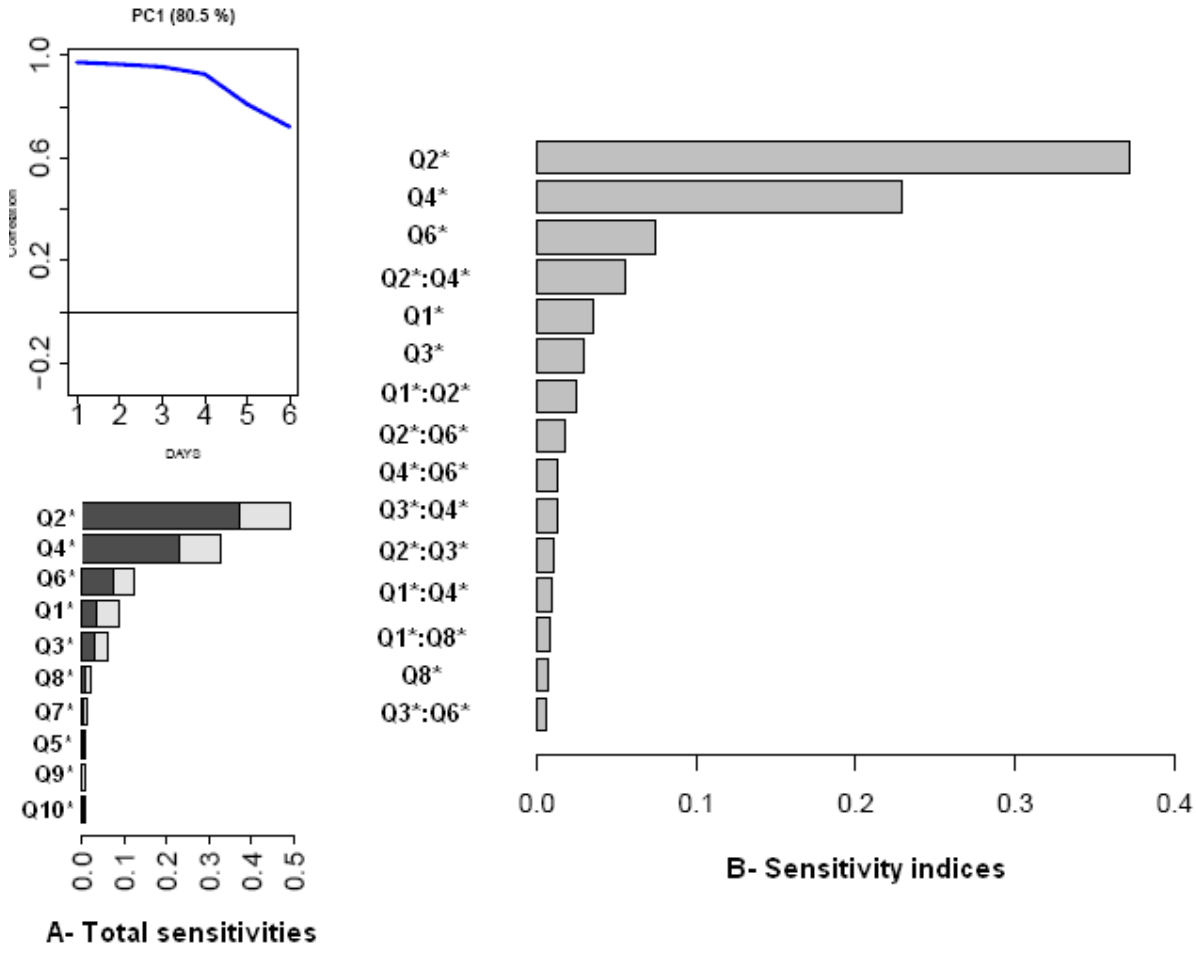

Figure 3.8. Sensitivity analysis of the means of six of the outputs (all except the abortion number) for the last simulation time point (week 260): results of the ANOVA performed for the first principal component (inertia: 80.5%). (A) Loadings defining the principal component for each time variable (in abscissa) and total sensitivities for the 10 probability distributions Q ranked in descending order. Sensitivities are split in main effect (black) and two-factor interactions (grey). (B) Sensitivity indices of the 15 main factorial terms (main effects or interactions) in descending order.

Table 3.5. Results of the third experiment of the sensitivity analysis: for the mean and standard deviation of each output, the proportion of total inertia represented by the first principal component and the corresponding most sensitive factors are given. Factors with total sensitivities ≥ 10%: one cross, factors with total sensitivities ≥ 25%: two crosses, factors with total sensitivities ≥ 50%: three crosses.

	Environment $E_{building}$		Environment E_{dry}		Prevalence of milk shedders		Prevalence of m/f shedders		Prevalence of milk shedders in a persistent way		Seroprevalence		Abortion rate		Joined analysis time point 260	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
inertia PC1	96.8%	90.7%	84.8%	83.5%	91.2%	74.4%	87.7%	76.4%	86.6%	57.1%	96.0%	74.7%	82.0%	57.6%	80.5%	56.6%
*Q1**					+	+	+	+	+	+	+	+	+	+		+
*Q2**	+++	+++	+++	+++	+++	+	+++	+	+++	++	++	+	+++	+	++	+
*Q3**																
*Q4**		+	++	++	++	+	+	+	+	++	+	+	+	+	++	+
*Q5**																
*Q6**					+	+	+	+	+	+	+	++	+	+	+	+
*Q7**																
*Q8**						+		+		+				+		+
*Q9**																
*Q10**																

The second and main part of our study consisted of a sensitivity analysis. This approach aims at improving the understanding of complex systems such as stochastic epidemiologic models with a view to suggesting possible targeted control strategies in livestock infections [43, 93, 156] or to assessing the effect of varying the input parameters on the economic impact associated with the disease [28, 165].

To perform sensitivity analyses, we used complete and fractional factorial designs. Alternative approaches are available (see for example [140] or [113]), but factorial designs are very convenient to control which main effects and interactions can be estimated and to manage a mixture of qualitative and quantitative factors. In the present study, some factors (α, β, β_{calv}, γ, γ_{calv}, *Q1, Q2, Q3, Q4* and *Q5)* were not scalars but probability distributions. Performing a sensitivity analysis with such types of factors is to our knowledge little known. We used multinomial distributions with three classes and defined the modalities of each such factor as four alternative sets of multinomial probabilities. This choice allowed flexibility as well as a fine control in the probability distributions that were simulated.

To cope with the dynamic and multivariate outputs of the model, the PCA+ANOVA approach [82] offered additional insight compared to single sensitivity analyses. Multivariate decomposition methods other than PCA could be used, but the key idea is that sensitivity analyses are now performed on synthetic and meaningful output variables. As the infection dynamics is composed of two phases (see on Figure 3.5 an initial phase of rapid evolution followed by a kind of steady state), the parameters influencing the dynamics may not be the same between the phases. We additionally performed a preliminary sensitivity analysis on the very initial phase of infection (first 26 weeks; results not shown). The most sensitive factors during the first 15 weeks were highly influenced by the initial conditions but very fast afterwards, the same factors as in experiments one or two were identified as the most influential ones. We then chose to conduct our study on the whole simulation period.

Another aspect concerns the stochasticity of our model which generates complex dynamics. Hence, attention has to be paid to this point in order to appropriately apply sensitivity analysis approaches that are mostly developed for deterministic models. Here, to be able to study with confidence how the variation in the outputs of our model can be apportioned to the uncertainty of epidemiological factors, we checked that the variability due to the model stochasticity (i.e. the within scenario variability) was negligible compared to the variability due to the input parameters variability (i.e. the between scenarios variability). More specifically, to provide reliable analysis on trends, means were calculated on 30 repetitions.

Standard deviations were also considered as they can provide complementary information on the most influential factors.

In summary, the method used has the major advantage of allowing to consider temporal varying outputs and thus to identify parameters influencing the dynamics over the 5 year simulation period. Moreover, it allows properly dealing with the model stochasticity.

According to our findings (first and second experiments), we can classify the eight parameters that have most influence on the *C. burnetii* infection dynamics in three categories, depending on the dynamics aspects they are involved in.

The first group, comprising the parameters related to the transition between health states, slightly influences the different prevalences and the abortion number. *q* (transition probability from *I-* to *I+*) is a physiological parameter associated to seroconversion and it was estimated based on data from a follow-up of five chronically infected herds [35]. However, we can assume that those five herds do not cover the whole potential range of variability of this factor, especially at the beginning of the infection when this parameter probably takes a different value depending on how recently the infection occurred in the herd. Further experimental or survey studies focusing on the start of the infection dynamics are needed to improve the knowledge of this parameter. The parameter *s* (transition probability from *C+* to *I+*), representing the intermittency of shedding, was inferred from data in the same study [35] and the estimated values were herd-dependent. It is biologically plausible to assume that a control measure such as vaccination could decrease this parameter and then have an impact on the prevalence of shedders. However, since, to our knowledge, no data is currently available, further studies are needed. The transition probability from $I^{+\,milk\,pers}$ to *C+*, r_2, and the proportion of cows going from *I-* to *I+* and becoming $I+^{milk\,pers}$, *pIp*, have an impact on the variability in the prevalence of persistent milk shedders. These parameters were not estimated from data, but calibrated so that simulated trajectories of prevalence of persistent milk shedders are consistent with field observations. Indeed, following Guatteo et al. [59], almost one milk shedder out of three was detected as persistent shedder over three months. In our case, the mean prevalence of milk shedders at time point 260 weeks is 3.3 times the mean prevalence of persistent milk shedders. While *pIp* seems difficult to be decreased, r_2, which rules the shedding duration of persistent milk shedders, could probably be modulated by control strategies. Although according to Astobiza et al. [16], an oxytetracycline treatment would not limit the duration of bacterial excretion in a dairy sheep flock, further studies are needed in order to determine if vaccination could decrease the length of the shedding period.

The second group of parameters that influenced the infection dynamics most is related to the characteristics of *C. burnetii* in the environment. In fact, ρ^{mf} (proportion of bacteria shed through mucus/faeces filling the environment compartment) and μ (mortality rate of *C. burnetii*) have a strong impact on the environmental bacterial load and to a lesser extent on the abortion number. Concerning the parameter ρ^{mf}, it is very difficult to quantify in practice which proportion of bacteria contained in milk, vaginal mucus or faeces contaminates the environment. Thus, we calibrated this parameter to match the environmental bacterial load estimated by Courcoul et al. [35]: the posterior medians for the environmental bacterial loads of each of the five chronically infected herds were comprised between 0.044 [0.005–0.143 for the 95% credible interval (CI)] and 0.558 units of environment [0.201–1.278 for the 95% CI]. Since those herds did not exhibit any clinical signs, we assumed that the simulated environmental bacterial load in herds with abortions should be slightly higher. The median of $E_{building}$ at time step 260 weeks is then 0.56 (with the percentiles 2.5^{th} and 97.5^{th} equal to 0.00 and 1.44 respectively). Concerning the parameter μ, given that *C. burnetii* withstands harsh environmental conditions [103], its life expectancy ($1/\mu$) within the farm in an infectious form was assumed to be five weeks in the standard scenario and two or 13 weeks (two extreme values) in the sensitivity analysis. However, this assumption is difficult to verify. Based on Dutch studies, it seems that within a month, more than 75% of the manure does not contain viable *C. burnetii* anymore but that the bacterium survives only a few days in the outer layer of the manure [161]. It is then difficult to calibrate the parameter μ which represents both the natural mortality of the bacterium and its removal due to the periodic cleaning of the cattle housing carried out by the farmer. However, it seems possible to influence μ by implementing environmental control measures such as increased cleaning of the farm.

Lastly, as suggested by the results of the first experiment (Table 3.4) and detailed in the third experiment (Table 3.5), the parameters which have the greatest impact on the infection dynamics are those governing the shedding levels (through their associated probability distributions Q^*), especially in mucus/faeces.

As shedding in mucus/faeces much more contaminates the environment as shedding in milk, it could seem surprising that the parameters governing the probability distribution of the shedding routes (α, β and γ) do not appear to influence the model outputs. This could partly be explained by the parameter values used, especially for *Qty*. The quantity of bacteria shed by a high level shedding cow is so high (compared to mid or low levels), that the probability distributions governing the levels (such as Q in experiments one and two) are by construction more important than those related to the shedding routes. However, although exact values of

parameters *Qty* are unknown, the standard values used in this study were calibrated to field data, which tend to support our findings.

This third experiment was of high interest as it allows highlighting the importance of a sub-category of animals. Indeed, the factor *Q2** (corresponding to the distribution of the levels of bacteria shed in mucus/faeces by seronegative shedders) has a strong impact on all the outputs, including the shedder prevalences and the environmental bacterial loads. To a lesser extent, the factors *Q4** and *Q6** (corresponding the distributions of the shedding levels in mucus/faeces of seropositive shedders) also have a significant impact on the variability of the shedder prevalences. The predominance of *Q2** over *Q4** and *Q6** can be explained by the larger simulated number of seronegative shedders than seropositive ones. In fact, due to the low standard value of *q* (the transition probability from *I-* to *I+*) and the high standard value of *m* (the transition probability from *I-* to *S*), the number of seropositive cows is very low during the three first years of simulation compared to the number of seronegative cows. The standard values of parameters *q* and *m*, which were estimated in chronically infected herds, are perhaps not perfectly appropriate to describe the initial phase of an infection and then could lead to overestimation of the influence of seronegative shedders. As suggested by Matthews et al. for *Escherichia coli O157* [105], identifying factors such as age, genetics, reproductive status or other management factors that might predispose an animal to high levels of shedding would be of undisputable interest. Moreover, control measures should aim at reducing the proportion of high shedders in mucus/faeces, such as phase I vaccines seem to do. According to Arricau-Bouvery et al. [13], vaccination dramatically reduced both abortions and excretion of bacteria in the milk, vaginal mucus and faeces. In Rousset et al. [136], the vaccine was effective in reducing massive bacterial shedding from a heavily infected goat herd. In the same way, Hogerwerf et al. [64] found that in uterine fluid, vaginal mucus and milk of vaccinated dairy goats, both the prevalence of shedders as well as the concentration of bacteria were reduced.

7. Conclusion

This work led to the identification of key parameters in *C. burnetii* infection dynamics based on an original model describing the bacterium spread within a dairy cattle herd composed of animals with heterogeneous shedding characteristics. The most influential parameters are associated to the probabilities governing the levels of shedding, especially for mucus/faeces shedders and to the characteristics of the bacterium in the environment. Some physiological parameters related to the intermittency of shedding or to the transition from one type of

shedder to another one also play a non negligible role. Our study also highlights parameters that have to be precisely assessed and then further investigated to improve the model accuracy and the understanding of the infection spread. Besides, interventions impacting those key parameters would be of great interest. The model developed here can be further used to assess over a longer time scale the effectiveness of different control strategies for *C. burnetii* infection, such as vaccination.

8. Acknowledgements

The authors would like to thank Alain Joly and Raphaël Guatteo from the Oniris-INRA group 'Bioaggression, Epidemiology and Risk Analysis' for fruitful discussions on the epidemiological model and on the way to take into account the heterogeneity of shedding. The collaboration between the Faculty of Veterinary Medicine of Utrecht and the French National Institute for Agricultural Research (INRA) was financially supported by the Netherlands Organisation for Scientific Research (NWO) and the French Ministry of Foreign and European Affairs through the Van Gogh Programme.

CHAPTER 4

ASSESSMENT OF THE COMPARATIVE EFFECTIVENESS OF THREE VACCINATION STRATEGIES IN *C. BURNETII* INFECTION

Picture : A. Senkowski

I- Why and how to include vaccination in models?

One of the main objectives of epidemiological modelling is to help public health or animal health decision makers designing guidelines for the management of infectious diseases by assessing the effectiveness of different control strategies. Various interventions are possible and they can be distinguished with respect to their aim. If they aim at preventing contacts between infected and susceptible individuals, contact tracing and quarantine can be implemented for humans as well as movement ban, market ban, and culling for animals. When the decrease in the susceptibility of susceptible individuals and/or the infectiousness of infected ones is aimed at, vaccination and prophylactic and/or therapeutic use of medications are appropriate measures. Vaccination, which is the topic of this chapter, does not aim to cure infected individuals. It consists in a preventive immunization: vaccine contains antigens which hopefully induce an immune response in the host, intended to be similar to the consequences of an infection. They are assumed to reduce the intensity of clinical signs, and/or to protect from the infection, and/or to stop, or at least to limit, the infectiousness when vaccinated individuals get infected.

In the next subsection, we will briefly describe different types of vaccination strategies that can be implemented to control human and animal diseases. In subsection 2, an example of modelling-based study aimed at vaccine effectiveness assessment will be presented. Finally, we will focus in section II on our study on the *C. burnetii* spread and the assessment by simulation of the effectiveness of different vaccination strategies in already infected cattle herds. This last part will be presented as it was submitted to the journal Veterinary Research.

1. Different types of vaccination programmes and their representation in epidemic modelling

For human diseases, pediatric vaccination is usually used to reduce the prevalence of endemic diseases like measles, polio or rubella. In a *SIR* model, it leads to consider that a fraction of newborns is effectively vaccinated and directly arrives into the *R* health state. The modeller has also to take into account that vaccines are often not fully efficient and that immunity is of limited duration. Therefore, individuals can need boosting. Pediatric vaccination is a long-term strategy and does not instantly lead to eradication of the infection [77]. However, a major advantage is that not all individuals need to be vaccinated to eradicate the infection. The reproduction ratio in a vaccinated population R' is equal to $(1-p)R_0$ with p, the fraction of

newborns vaccinated and R_0 the basic reproduction number of the studied infection. Thus, to get $R' < 1$, the proportion of vaccinated newborns p has to be at least equal to $1-1/R_0$. This phenomenon is called "herd effect": the vaccination of a portion of the population provides protection to unprotected individuals [70].

For non endemic infections, a mass-vaccination program can be implemented when there is an increased risk of epidemic. In this case, there is a race between the infection spread and the vaccination programme and the best way to control the epidemic is a strong and early response [77]. The consequence of mass-vaccination in terms of modelling is often the addition of a new health state for vaccinated individuals. Mass-vaccination programmes were well-studied for the 2009 *Influenza* H1N1 pandemic. As an example, Sharomi et al. [143] showed that in Canada, with the estimated vaccine efficacy of 80%, at least 60% of Manitobans needed to be vaccinated in order to effectively control the pandemic and that the timely implementation of the mass vaccination program was crucial.

It is also possible to first protect individuals that are most at risk. This strategy is called "targeted vaccination" and induces the representation in the model of host risk categories. In human diseases, this strategy was used in France during the *Influenza* H1N1 vaccination campaign, where people at risk of developing complications (e.g. pregnant women, young children, immunodepressed people, people with chronic broncho-pulmonary affections, etc…) were first vaccinated. For animal diseases, vaccination of animals in contact with infectious ones or ring vaccination around confirmed cases are other examples of targeted vaccination programmes (e.g. the FMD model of Keeling et al. [75]).

For human diseases, pulse vaccination can also be implemented: it consists in periodic vaccination of certain age cohorts [77]. The aim is to maintain the proportion of susceptible individuals below the threshold enabling the infection to spread. This type of programme is then composed of two stages: a punctual vaccination of a high proportion of children of a given range of age, followed by a period (some years) without vaccination. Such programmes are logistically easier to implement than continual pediatric vaccination. They can be used for childhood diseases like measles [4]. Models are a useful mean to determine the maximum permitted interval between pulses function of the epidemiological, demographic and vaccination factors [117].

In animal diseases, vaccination is sometimes implemented in an already infected herd. The aim of such a programme can be preventive only: the immunization of the still susceptible individuals can decrease the probability of becoming infected, the intensity of their clinical

signs and infectiousness if they become infected. For some infectious diseases like the Infectious Bovine Rhinotracheitis, vaccinating the already infected animals can also limit their clinical signs and infectiousness and therefore be an effective measure to limit the pathogen spread [137]. However, as vaccines do not have any curative effect, vaccination is first a preventive strategy.

At last, it has to be highlighted that, when analyzing the results of a simulation model dealing with vaccination, decision makers have to keep in mind logistical as well as social and economic limitations (e.g. number of vaccine doses available, time required to vaccinate an individual, non observance of the vaccine recommendation, etc.).

2. An example of model aimed at assessing the effectiveness of vaccination

Until recently, models assessing the effectiveness of vaccination were rather little used in animal health. They mainly dealt with FMD [75] and rabies [146]. In the recent years, they have been developed. They were used for example to assess the potential impact of imperfect *Salmonella* vaccines on the prevalence of infection in infected dairy herds [92], to identify key factors influencing the apparent success of vaccination to control Bluetongue virus Serotype 8 spread in Great Britain [148], and to evaluate different vaccination strategies against brucellosis in bison [155] and against BVDV in cattle [144].

In order to illustrate in more details the interest of using modelling-based study when exploring vaccination strategies, we are now going to present the model of Suppo et al. [146] aiming at assessing the effectiveness of two prophylactic methods (contraception and vaccination) for rabies control in fox populations. In Europe, fox populations tend to increase, which could impede the success of oral-vaccination campaigns because of the growing number of susceptible animals. The objective of this study was to determine the potential interest of fertility control through the use of baits filled with a contraceptive vaccine in conjunction with a rabies vaccine. The model was compartmental, deterministic, and in discrete-time. It took into account host heterogeneity: the fox population was structured in age, sex, and health state (Figure 4.1). Besides, it was spatially-explicit: a rectangular domain was divided into cells corresponding to the size of an average fox's home range. Demography was simulated either exponentially increasing or density dependant through survival and birth rates. Dispersion of young foxes was represented: these animals had the possibility to settle in a new cell, the probability of settling in a given cell being function of the distance between the former and the new cell. The transmission of the infection could occur between foxes living in the same or

adjacent cells (this transmission was function of β, the transmission rate from an infectious to a susceptible fox), and during dispersal, when infected young foxes reached a new cell.

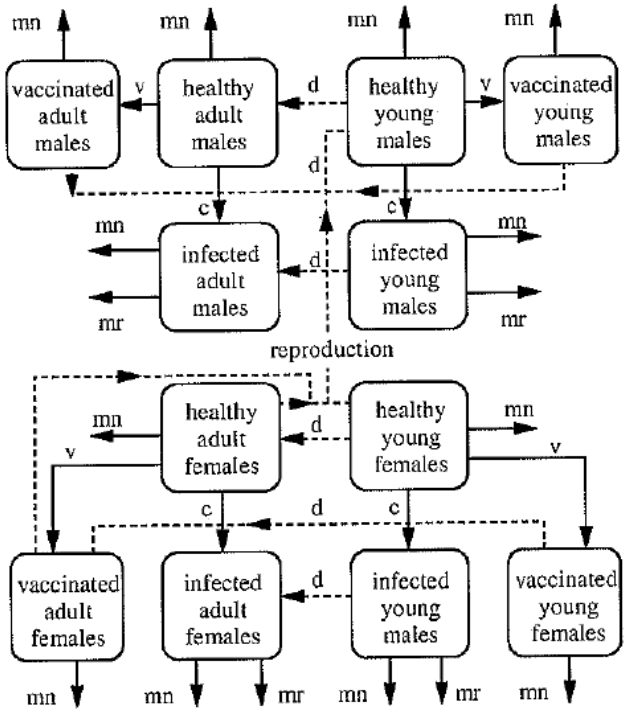

Figure 4.1. Interaction between the 12 classes of foxes. mn, natural mortality; mr, mortality induced by rabies; v, vaccination; d, dispersal; c, contamination
From Suppo et al. [146]

Two vaccination campaigns per year were simulated as well as fertility-control campaigns, occurring once a year and being effective on females during one breading season. Vaccination was modelled by adding new compartments for vaccinated animals. Besides, in case of fertility control, births were decreased by (1-*st*), with *st* the sterilization rate. The initial conditions consisted in the introduction of a pair of exposed adult foxes in a single cell located at the centre of the domain. Each of the other cells contained a pair of healthy foxes. The effectiveness of contraception and vaccination was evaluated for different values of birth and transmission rates and for each of the type of population growth (exponentially increasing or density dependant). In exponentially increasing populations (which seems currently the case in European countries), for 4 to 7 cubs per litter per female, the control of the infection with the sterilization programme alone was impossible because the healthy population would go extinct before rabies was eradicated. For a programme with only vaccination, the vaccination rate had to be high, especially for low transmission rates and high births rates, to lead to rabies extinction. In the field, a maximum of 70% of the population can be vaccinated during

vaccination campaigns. There were then cases for which vaccination alone failed to eradicate the infection. However, a combination of both fertility control and vaccination decreased the birth rate to a value requiring a lower vaccination rate and was then effective to eradicate the virus. Thus, those results suggested that contraception could be a possible additional method to control rabies outbreaks in highly dense fox populations. Nevertheless, the authors highlighted the need of further studies including fox culling, changes in spacing strategies when fox density increases, and the influence of dispersal in the recovery of healthy populations, to draw robust conclusions.

In the model developed in this example, it was possible to simulate the same control programmes (i) in both an exponentially increasing and a density-dependant host population growth, and (ii) for different values of birth and transmission rates. The conclusions regarding the effectiveness of control measures were different regarding the characteristics of the fox population and virus transmission. Therefore, models are a useful tool to assess the effectiveness of control strategies in different epidemiological situations.

II- Manuscript: Modelling effectiveness of herd level vaccination against Q fever in dairy cattle

Aurélie Courcoul[1,2], Lenny Hogerwerf[3], Don Klinkenberg[3], Mirjam Nielen[3], Elisabeta Vergu[4], François Beaudeau[1,2,5]

[1]INRA, UMR1300 Bio-agression, Epidémiologie et Analyse de Risque, Nantes, France
[2]Oniris, UMR1300 Bio-agression, Epidémiologie et Analyse de Risque, Nantes, France
[3]Utrecht University, Faculty of Veterinary Medicine, Utrecht, the Netherlands
[4]INRA, UR341 Mathématiques et Informatique Appliquées, Jouy-en-Josas, France
[5]Université Nantes, Angers, Le Mans, France

Submitted to Veterinary Research

1. Abstract

Q fever is a worldwide zoonosis caused by *Coxiella burnetii* which induces reproductive disorders in livestock. Ruminants are recognized as the most important source of human infection. The control of this infection in cattle is crucial to limit both the infection in livestock and the zoonotic risk. Although vaccination is currently advised in the field, the comparative relevance of different vaccination protocols in terms of the duration of the vaccination campaign and category of animals to be targeted has never been explored. Our objective was to compare, by simulation, the effectiveness of three different vaccination strategies in an already infected dairy cattle herd.

We used a stochastic individual-based epidemic model coupled with a model of herd demography to simulate three temporal outputs of shedders prevalence, environmental bacterial load and number of abortions and to calculate the infection extinction rate. For all scenarios, the temporal outputs strongly decreased with time at least in the first years of vaccination. However, vaccinating only three years is inadequate to stabilize these dynamic outputs at a low level. Vaccination of both cows and heifers is more effective than vaccinating heifers only. For heifers only, (i) the outputs decreased much slower and never reached the effectiveness of full herd vaccination, (ii) the infection extinction rate is twice as low as well.

Besides valuable indications on vaccination effectiveness, our model could also be adapted in further studies to simulate and asses other Q fever control strategies such as environmental and hygienic measures.

2. Introduction

Q fever is a zoonotic disease caused by *Coxiella burnetii*, a bacterium found worldwide in a wide range of animals. In ruminants, the infection may cause abortions, infertility, metritis or chronic mastitis [5, 20, 26, 125], which can lead to non negligible economic losses for the infected herds. Furthermore, since 2007, Q fever has become an important public health problem in several parts of Europe [72, 121, 159]. Although Q fever is asymptomatic in 60% of human cases, it can lead to acute or chronic infections and cause flu-like syndrome, hepatitis, pneumonia, endocarditis or abortions[12]. In the Netherlands, where a steep increase in the number of human cases was observed in 2007, 2008, and 2009, a link has been established between some human cases and farms of small ruminants where abortions due to Q fever were detected [141]. Ruminants are indeed recognized as the main source of human infection [54, 109]. Infected animals shed large quantities of bacteria into the environment through faeces, vaginal mucus, urine, milk and especially parturition products [11, 20, 59]. *C. burnetii* survives very well in the environment and contaminates aerosols and dust [167]. These infected particles are the main route of infection for both animals and humans. Due to its importance in both animal and public health, the control of this infection is crucial. Therefore, any control measure leading to a decrease in the prevalence of shedders and in the environmental bacterial load seems a key point to limit both the spread of the infection in ruminants and the zoonotic risk.

Nowadays, in infected cattle herds in France, control measures against Q fever consist of environmental measures such as destruction of placentas or disinfection of births locations, antibiotic treatment like oxytetracycline injections during the last month of gestation, and vaccination [132]. Observations concerning antibiotics are contradictory. In Berri et al. [22], antibiotics in sheep suppressed in the long run both the abortions and the shedding of *C. burnetii,* whereas in Astobiza et al. [16], the oxytetracycline treatment neither prevented the shedding of bacteria nor limited the duration of bacterial excretion. The EFSA concluded that, as antibiotic treatment in animals is not effective in influencing the epidemiology of infection, and as widespread antibiotic usage is inadvisable because of the development of resistance, antibiotic treatment for *C. burnetii* infections should be avoided [39]. According to Rodolakis

[12] ECDC, Risk assessment on Q fever, (2010) [on line] http://ecdc.europa.eu/en/publications/Publications/1005_TER_Risk_Assessment_Qfever.pdf [consulted 25 August 2010]

et al. [132], vaccination would be an efficient tool to control the disease. Vaccination with a phase I vaccine in cattle was shown to suppress the shedding in milk, placenta, uterine fluid, vagina and colostrum [25, 139]. More recently, Arricau-Bouvery et al. [13] compared the efficiency of phase I and phase II vaccines in goats: the phase I vaccine prevented abortions and dramatically reduced the frequency of bacterial shedding in the milk, vaginal mucus and faeces, while the phase II vaccine did not affect the course of infection. In Rousset et al. [136], the vaccine appeared neither able to prevent infection in exposed kids, nor to clear infection in infected goats, but effectively reduced the level of shedding in a heavily infected herd. Hogerwerf et al.[13] also found that both the prevalence of shedders and the bacterial load in uterine fluid, vaginal swabs and milk were reduced in vaccinated dairy goats. Besides, according to Guatteo et al. [61], susceptible cattle that were vaccinated when non pregnant had a five times lower probability to become a shedder than an animal receiving placebo.

Thus, in the field, vaccination is often recommended in infected herds after the occurrence of abortions due to Q fever. However, the studies assessing the vaccination efficacy in ruminants were carried out in experimental conditions or for a limited period of time and they evaluated the effect of the vaccine mostly at the individual level. Therefore, it is difficult to extrapolate those results to the case of a whole herd vaccination over several years. Another point to consider is that vaccination generally takes place in the field in infected herds without any preliminary individual diagnostic tests. Some cows may be vaccinated when still susceptible while others are already infected. Further studies are needed to assess the overall effectiveness of such vaccination programmes in cattle herds. Different vaccination strategies can be implemented: the duration of the vaccination programme as well as the category of vaccinated animals (e.g., the whole herd or the heifers only) have to be determined. To assess the long run effectiveness of these different strategies in reducing the infection prevalence or the environmental bacterial load, field studies are not optimal: no reference situation (without control strategy) is generally available, and long-term observations must be performed, making these studies very costly and even unfeasible. Modelling is therefore a convenient approach as it provides means to compare the effectiveness of different potential management strategies [77]. The use of mathematical models is nowadays widely used to compare control measures for both human [1, 143, 160] and animal infectious diseases [7, 17, 49, 87]. For *C. burnetii* infections, it would allow testing a wide range of vaccination strategies in different initial situations.

[13] Hogerwerf L., Van den Brom R., Roest H.J., Bouma A., Vellema P., Pieterse M., Dercksen D., Nielen M., Vaccination of dairy goat herds reduces *Coxiella burnetii* prevalence and bacterial load in goat excret. *Submitted for publication* in Emerging Infectious Diseases.

The objective of this model study is to assess the comparative effectiveness of several vaccination strategies in an already infected dairy cattle herd. The criteria considered for efficacy evaluation were changes in the prevalence of shedders, the environmental bacterial load, the number of abortions, as well as in the extinction rate of infection.

3. Materials and methods

A model representing the *C. burnetii* infection dynamics in a standard French dairy cattle herd and different vaccination strategies was elaborated based on a previous variant model not including interventions. First of all, the epidemic model representing the natural course of infection (i.e. without any control strategy) will be briefly described, then the inclusion of vaccination will be presented and finally, the different vaccination scenarios that we tested will be explained in detail.

a. General description of the epidemic model of the natural course of infection

The model represents the spread of the bacterium in a dairy herd of lactating and dry cows (diagram flow in Figure 4.2 and parameters in Table 4.2 of subsection 7. Supplementary material). It is a stochastic individual-based model in discrete time with a time step of one week. Each cow is in one of the six mutually exclusive health states at a given time: S (non-shedder apparently susceptible cow), I_1 (shedder which has the possibility to eliminate the bacterium and become S again), I_2 (shedder which does not have anymore the possibility to become S again), I_3 (shedder which does not have anymore the possibility to become S again and which sheds in milk at higher levels and for a longer period of time than I_2 – health state described in Guatteo et al. [59]), C_1 (non-shedder but still infected cow), C_2 (non-shedder which was C_1 in the past but eliminated the bacterium). Moreover, as a great heterogeneity between *C. burnetii* shedders has been described [11, 37, 59, 131], this individual variability in the shedding routes and the shedding levels (i.e. the quantities of bacteria shed) is taken into account in the model. Sub-categories are then defined for the shedder cows with respect to the shedding route. Thus, an I_1 or I_2 cow can shed in (1) milk only (denoted by I_1^m or I_2^m respectively), (2) vaginal mucus and/or faeces (I_1^{mf} or I_2^{mf} respectively), or (3) milk and either vaginal mucus or faeces or both (I_1^{mmf} or I_2^{mmf} respectively). In the same way, an I_3^m sheds in milk only and an I_3^{mmf} sheds in milk and vaginal mucus and/or faeces (by definition, an I_3 animal always sheds in milk and can not be in the I_3^{mf} state).

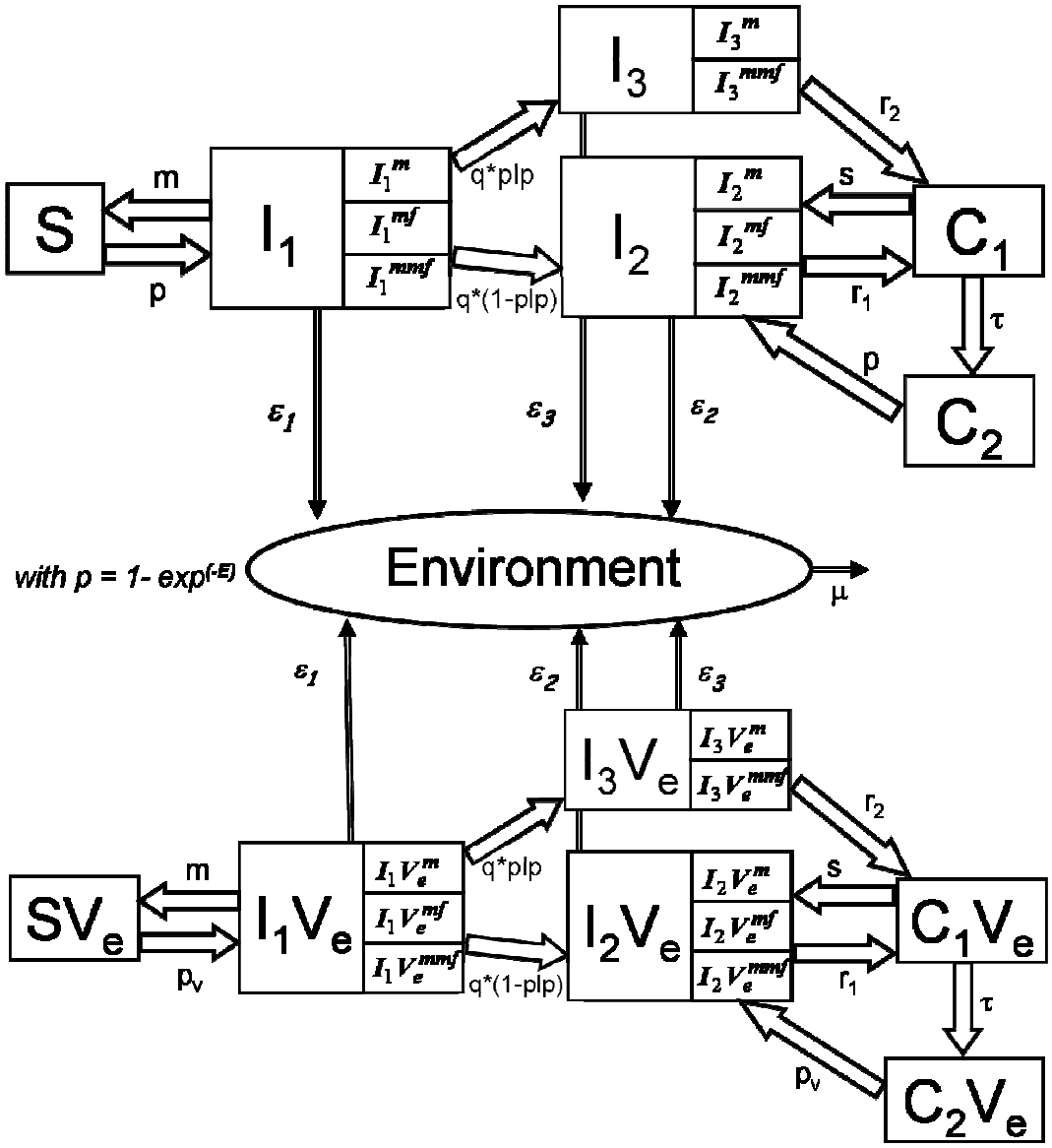

Figure 4.2. Flow diagram describing the modelled spread of *C. burnetii* within a vaccinated cattle herd. The health states are: S, non-shedder apparently susceptible cow, I_1, shedder which still has the possibility to eliminate the bacterium and to become S again, I_2, shedder which does not have anymore the possibility to become S again, I_3, shedder which does not have anymore the possibility to become S again and sheds in milk in a persistent way, C_1, non-shedder but still infected individual and C_2, non-shedder which was C_1 in the past but eliminated the bacterium. The V_e states (SV_e, I_1V_e, I_2V_e, I_3V_e, C_1V_e and C_2V_e) are defined in the same way as S, I_1, I_2, I_3, C_1 and C_2 respectively, except that these animals have been vaccinated when susceptible and non pregnant and are then assumed "vaccinated in an effective way" (V_e). I and IV_e cows are in the subcategory m if they shed in milk only, mf if they shed in vaginal mucus/faeces only and mmf if they shed in milk and vaginal mucus/faeces. E represents the environmental bacterial load and p, the probability of infection or reinfection for non V_e individuals, is equal to $1-\exp^{(-E_t)}$. p_v is the probability of infection or reinfection for V_e individuals, which is a fraction of p. The other model parameters are presented Table 4.3. of subsection 7. Supplementary material. ε_1, ε_2, ε_3, ε_1V_e, ε_2V_e *and* ε_3V_e are the quantities of bacteria shed during a time step by an individual I_1, I_2, I_3, I_1V_e, I_2V_e and I_3V_e respectively and contaminating the environment. For any shedder, ε represents the sum, for each shedding route, of the quantity of bacteria released, Qty, times ρ its fraction reaching the environment of the herd.

The possible transitions between health states are represented in Figure 4.2. Shedders (I_1, I_2 and I_3) fill the environment compartment (E) with bacteria: the quantity of bacteria arriving into the environment during a time step is the sum for all the shedders and all the shedding routes of the quantity of bacteria shed, *Qty* (the shedders can shed at low, moderate or high level, *Qty* being different for each of these levels according to probability distributions Q), times ρ the impact of this shedding on the environment (i.e. ρ is the fraction of bacteria shed which arrives into the environment of the herd - for more detail see Table 4.3 of subsection 7). The probability of infection or re-infection, p (transition from S to I_1 or from C_2 to I_2) is expressed at each time step as $1 - \exp^{(-E_t)}$ where E_t is the quantity of bacteria in the herd environment at time t (one unit of E_t corresponding to a probability of transition from S to I_1 of (1 - 1/e)). The mortality rate of *C. burnetii* in the environment, μ, includes the natural mortality of the bacterium and its removal in relation to the periodic cleaning of the cattle housing carried out by the farmer.

As abortions are the main clinical signs attributable to *C. burnetii* infections [132, 135], they are also represented in the model: a cow can abort after a transition from S to I_1, from C_1 to I_2 or from C_2 to I_2 but only once in her life. If the cow aborts in the first or second third of gestation, she sheds through the mucus/faeces a moderate quantity of bacteria *Qty,* whereas if the abortion occurs in the last third of gestation, a high quantity of bacteria is released through this shedding route.

The epidemic model was also coupled to a model of population dynamics in order to represent the gestation and lactation cycles of each cow. In short, for each cow the lactation number is represented, as well as the stage of lactation, the stage of gestation, the abortion history, the health state and the shedding characteristics (if the cow is shedding).

b. Representation of the vaccination

Based on Guatteo et al. [61], we assumed that the vaccine is effective when applied to non pregnant uninfected individuals. Thus, in the epidemic model, non pregnant S and C_2 individuals become partly protected when vaccinated and move into the 'vaccinated in an effective way' (V_e) states (bottom of Figure 4.2). Pregnant S and C_2, as well as all I_1, I_2 and C_1 are what we defined the uselessly vaccinated: the vaccine has no effect on the infection dynamics in these animals, and they keep moving between the states S, I_1, I_2, I_3, C_1 and C_2 (top of Figure 4.2). Six additional health states are defined for the V_e individuals. SV_e and C_2V_e individuals can get infected and become I_1V_e or I_2V_e respectively with a decreased transition rate p_v (equal to a fraction of p). Except for this difference between p and p_v, the V_e animals can evolve through the same health states with identical transition rates as the non V_e animals.

Regarding the shedding levels, according to Guatteo et al. [61], the only quantified bacterium load of a V_e shedder was lower than the lowest bacterium load of the placebo cows. Besides, in Rousset et al. [136], the bacterial loads in vaginal swabs were lower in vaccinated than in non vaccinated animals. Therefore, we assumed that no high level shedding is possible for V_e animals and that the probability to shed at a low level is increased (expressed through probability distributions QV_e in Table 4.3 of subsection 7). Finally, based on Arricau-Bouvery et al. [13], it was assumed that the V_e cows cannot abort.

c. Vaccination scenarios

- ✓ **Scenario 1: vaccination over the whole simulation period (10 years)**

At the start of the simulation, all the cows are vaccinated and all the heifers entering the herd of cows are assumed to be SV_e (susceptible and vaccinated when non pregnant). In addition, all the animals are boosted every year: there is no loss of immunity and no possible transition from the V_e states to the non V_e states.

- ✓ **Scenario 2: vaccination for a limited period of time (3 years)**

The assumptions are the same as those of scenario 1 except for the vaccination duration. Here, the herd is supposed to be vaccinated for 3 years. At the end of this 3 year period, two assumptions regarding the evolution of immunity were explored.

  - Scenario 2A: immunity lasts for one year. One year after the end of the vaccination period, the V_e animals loose their immunity and move to the non V_e equivalent states (e.g. an I_2V_e cow becomes an I_2 cow).
  - Scenario 2B: lifelong immunity. After the vaccination period, the V_e animals do not loose their immunity and keep moving within the V_e states until the end of their life.

- ✓ **Scenario 3: vaccination of the heifers only over the whole simulation period (10 years)**

At the start of the simulation the cows are not vaccinated: they stay in the non V_e states and progress through infection states. Only the heifers arriving thereafter are assumed to be vaccinated. These animals are in the SV_e state when entering the dairy herd. In addition, they are boosted every year: there is no loss of immunity and no transition from the V_e states to the corresponding non V_e states.

- ✓ **Negative control**

No control programme is implemented and all the animals progress through the non V_e states.

d. Parameters and initial conditions

The values of epidemiologic parameters are displayed in Table 4.3 of subsection 7. Parameters m, q, r_1, s and μ were fixed at their values estimated through Bayesian inference using data from five French chronically infected dairy cattle herds [35]; probability distributions of shedding related parameters, α, β, β_{calv}, γ, γ_{calv}, (governing the partition in subcategories according to the shedding route) and $Q1$, $Q2$, $Q3$, $Q4$ and $Q5$ (characterizing the shedding levels), were qualitatively calibrated to match field data. The parameters governing the demography and herd management (Table 4.2 of subsection 7) were chosen to represent a standard French dairy cattle herd.

The transition rate p_v was parameterized using the probability for an initially susceptible animal to become a shedder in Guatteo et al. [61], which is equal to 0.21 with a 95% confidence interval of 0.05-0.90. Thus, we performed the simulations with p_v=0.21p. However, in scenario 1, two additional values were also tested (p_v=0.05p and p_v=0.90p) in order to determine the influence of this parameter value on the model output.

We simulated 100 repetitions of the introduction of a primiparous I_2 cow which has just calved into a fully susceptible herd of 50 cows to generate infected herds. We let the model run until three abortions had occurred during a period of 12 months to initiate reactive vaccination. This limit was motivated by the fact that testing for a large panel of abortive pathogens (including *C. burnetii*) is usually performed in France from the 3rd abortion within the calving period. Thus, we obtained 100 so called "initial herds", different from each other. Then, for each initial herd, the three vaccination scenarios and the negative control scenario were run once over a 10-year simulation period.

e. Outputs of the model

The mean prevalence of shedders, the number of abortions per herd per year and the environmental bacterial load were the model's dynamic outputs of interest. In addition, for each scenario, the rate of extinction over the 10 year simulation period was calculated as the ratio between the number of extinct trajectories and the total number of repetitions. The infection was assumed to be extinct when there were no more I, IV_e, C_2 and C_2V_e cows in the herd at the end of the simulation time.

Moreover, as the vaccine was reported to be effective for susceptible animals only [61], we tested whether the vaccination schedules were less effective when applied in heavily infected herds. Thus, the extinction rate was separately calculated in scenario 1 for several classes of initial prevalences of shedders, or initial prevalences of shedders in milk. The 100 simulated

herds were split by threshold prevalences at the 20th and 80th percentiles, resulting in the following classes: initial total shedders prevalence of [0-15%], [15%-40%] and [40%-100%] or initial milk shedders prevalence of [0-6%], [6%-20%] and [20%-100%].

4. Results

a. Description of the herds at the start the vaccination strategy

At the start of simulations, the mean prevalence of shedders (over 100 initial herds) is equal to 28.5% (min: 0.0%, max: 63.8%) and the mean prevalence of milk shedders amounts to 13.3% (min: 0.0%, max: 37.9%). In a herd, 92.8% of the cows on average have been shedders for at least one time step (min: 40.8%, max: 100%). The mean environmental bacterial load is 0.30 units (min: 0.02, max: 0.98) and the herds consist of 49.8 cows on average (min: 43, max: 58).

b. Influence of the vaccination scenarios on the temporal model outputs

If no control strategy is implemented, the mean prevalence of shedders, the mean environmental bacterial load and the mean number of abortions increase to a steady state of respectively 47%, 1 unit of environment and 4.1 abortions per herd per year. On the contrary, for any vaccination scenario, all these outputs decrease with time at least for the first years of vaccination (Figure 4.3.a, 4.3.b, 4.3.c). In scenario 1 (vaccination of heifers and cows during 10 years), the decrease covers the whole period. In scenario 3 (vaccination of heifers only for the whole simulation time), the decrease is much slower in the first three years of vaccination than in scenario 1: the latter allows reaching a mean prevalence of shedders of 5% and a mean environmental load of 0.05 respectively 2 and 1.5 years sooner than scenario 3. At the end of the vaccination period, the mean prevalence of shedders and environmental bacterial load are respectively equal to 2.8% and 0.04 units in scenario 1 and 5.0% and 0.06 units in scenario 3. The mean number of abortions in the first year of the vaccination program is equal to 2.5 and 3.6 in scenarios 1 and 3 respectively. In scenario 2, there is an increase in the mean prevalence of shedders, the yearly number of abortions and the environmental bacterial load, after the vaccination is ceased. For scenario 2A, this increase occurs immediately after the loss of immunity, whereas for scenario 2B (lifelong immunity), the increase is almost zero in the first year without vaccination and more progressive afterwards. Thus, the mean prevalence of the shedders is around 14% for both scenarios 2A and 2B three years after the simulation start and increases to respectively 45.4% and 32.0% eight years after the simulation start.

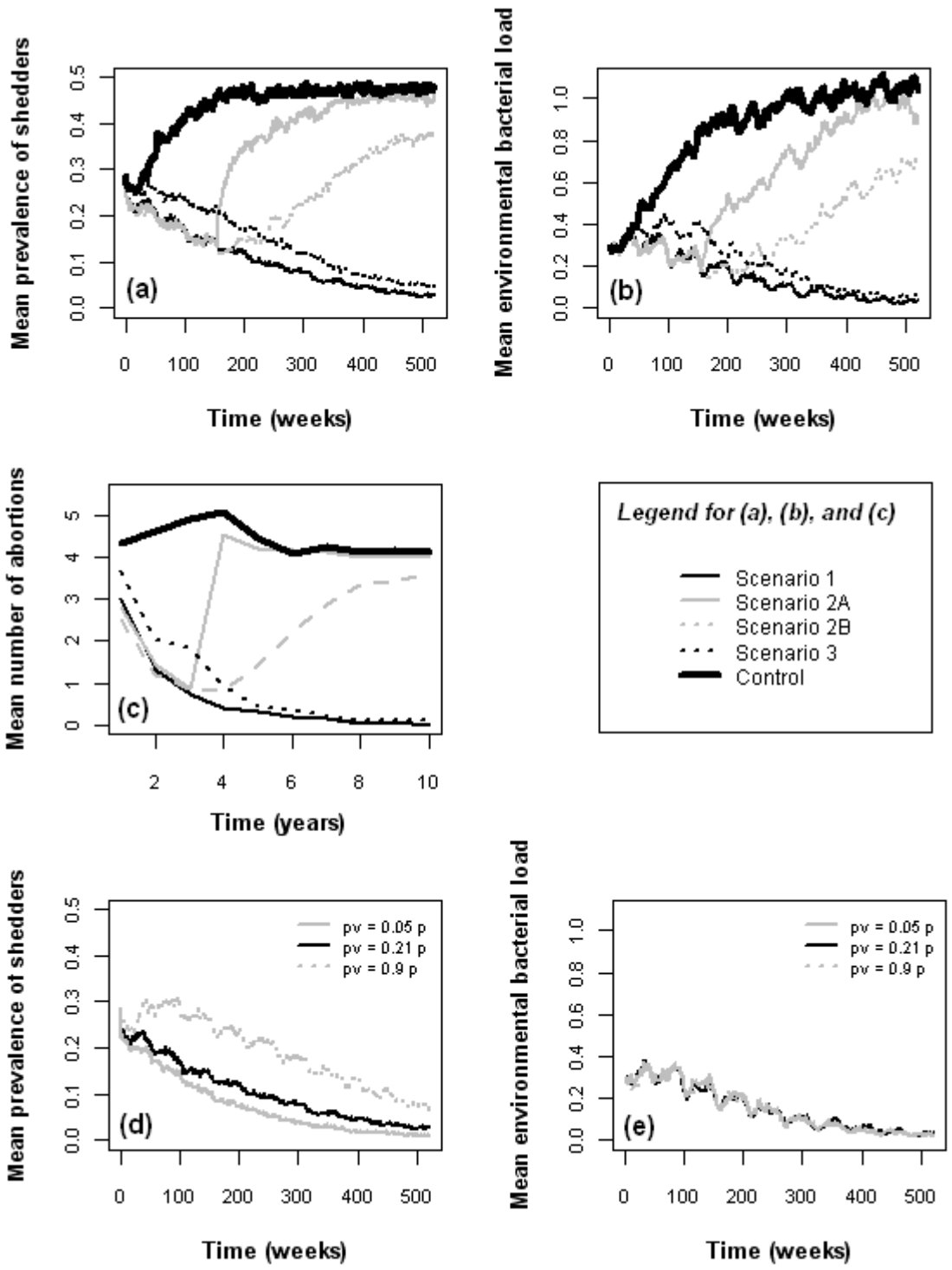

Figure 4.3. Temporal dynamics of the mean prevalence of shedders (a), the mean environmental bacterial load (b) and the mean number of abortions (c) with respect to the vaccination scenarios. Scenario 1: vaccination of heifers and cows for a 10-year period (black line); scenario 2: vaccination of heifers and cows for a 3-year period with (scenario 2A – grey line) or without (scenario 2B –grey dotted line) loss of immunity one year after at the last vaccination; scenario 3: vaccination of heifers for a 10-year period (black dotted line); control: no vaccination (black thick line). Temporal dynamics of the mean prevalence of shedders (d) and mean environmental bacterial load (e) in scenario 1 with different values of p_v (transition rate from SV_e to I_1V_e).

The mean number of yearly abortions increases from 0.9 the third year after the start of vaccination to 4.2 and 3.7 abortions per herd respectively during the eighth year after vaccination starts.

c. Influence of the p_v values on the model dynamics

As shown in Figure 4.3.d and 4.3.e for the scenario 1, the mean prevalence of shedders is highly influenced by the values of p_v, whereas the mean yearly number of abortions (results not shown) and the mean dynamics of environmental bacterial load are not affected by this parameter. For $p_v = 0.9p$, the mean prevalence of shedders is almost stable within the first three years of vaccination and decreases afterwards to reach 9.3% at the end of the simulation time. On the contrary, when considering $p_v = 0.05p$, the decrease is much faster and the mean prevalence of shedders is less than 1% at the end of the simulation time.

d. Influence of the vaccination scenarios and of the p_v values on the extinction rate

Whereas the extinction rate is nil when no control programme is implemented, it varies from 4% to 42% between the vaccination scenarios and the values of p_v (Table 4.1). It appears that most of the extinctions occur late: as shown on Figure 4.4 for the scenario 1, only one third of the extinctions happen in the six first years of the vaccination programme.

Table 4.1. Extinction rate and mean time to extinction for each of the vaccination scenarios. Control: no control programme; scenario 1: vaccination of heifers and cows for a 10-year period; scenario 2: vaccination of heifers and cows for a 3-year period with (scenario 2A) or without (scenario 2B) loss of immunity one year after at the last vaccination; scenario 3: vaccination of heifers for a 10-year period.

Criteria	Scenario						
	Control	1 $p_v = 0.05p$	1 $p_v = 0.21p$	1 $p_v = 0.9p$	2A	2B	3
Extinction rate	0.00	0.48	0.42	0.18	0.04	0.13	0.20
Mean time to extinction	-	week 349	week 361	week 275	week 84	week 216	week 411

The extinction rates for scenario 1 and $p_v = 0.21p$ are presented in Figure 4.5 according to the three classes of initial shedders prevalence and milk shedders prevalence. There is no significant difference between them (χ^2 tests, $p>0.05$). However, the extinction rates tend to be lower when the initial prevalences of shedders are high.

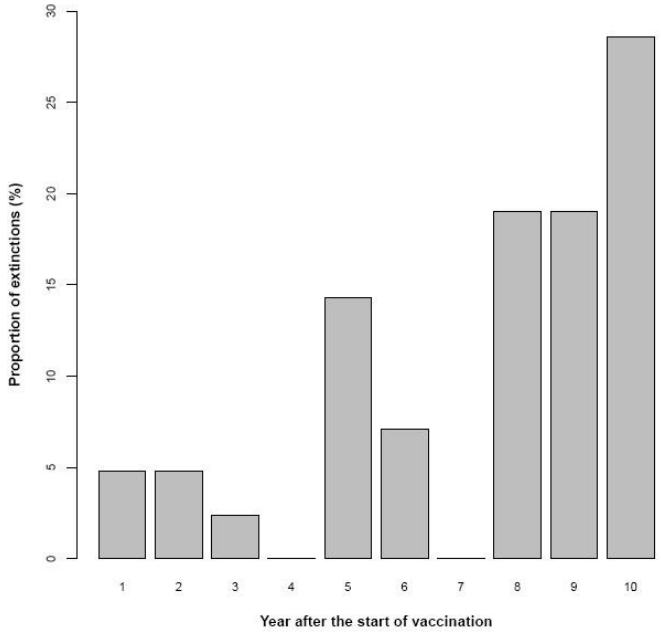

Figure 4.4. Proportion of extinctions (amongst the 42 extinct trajectories of scenario 1) according to the year after the start of vaccination when they occur.

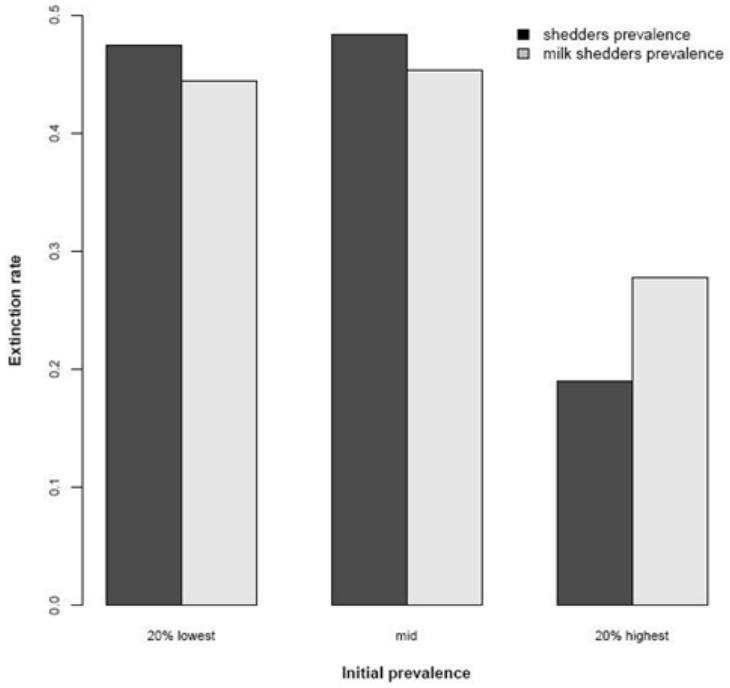

Figure 4.5. Extinction rate, for scenario 1 and $p_v = 0.21p$, stratified in 3 classes according to the initial prevalence of shedders (black bars) or milk shedders (grey bars) at the start of vaccination. Amongst the 100 runs, the 1st class comprises the trajectories with the 20% lowest initial prevalences and the 3rd class those with the 20% highest initial prevalences. The 2nd class includes the other repetitions.

5. Discussion

In this study, we modelled the long term effectiveness of three different vaccination strategies in an infected dairy cattle herd [(1) vaccination of the whole herd for 10 years, (2) vaccination of the whole herd for 3 years and (3) vaccination of the heifers only for 10 years] and showed that scenario 1 was the most effective control strategy. In fact, the three vaccination strategies all reduced the prevalence of shedders, the environmental bacterial load and the number of abortions. However, their effectivenesses are not equivalent. As the infection is most often not eradicated in the first years of vaccination, an early cessation of vaccination (scenario 2) will be ineffective in the long run. Its short-term effect on the infection dynamics depends on the lifetime of immunity for efficiently vaccinated cows. According to Rodolakis et al. [133], in infected herds, more than 80% of the vaccinated cows still had immune markers one year after vaccination. However, at the same time, less than 60% of the vaccinated heifers were still skin-test positive. In the field, this means that immunity should last between one year (scenario 2A) and life long (scenario 2B). In that context, the increase of the prevalence of shedders, the environmental bacterial load and the number of abortions should not be observable in the first months following the cessation of vaccination. Nevertheless, the infection is spreading again. Thus, before stopping a vaccination programme on a farm, it seems essential to determine the presence or absence of *C. burnetii* in the herd. Diagnostic tests at a herd level (e.g. PCR in bulk tank milk) can probably be helpful [58], although they are imperfect.

According to our simulations, when only the heifers are vaccinated yearly (scenario 3), the decrease in the prevalence of shedders, the environmental bacterial load and the number of abortions is much slower than when all the animals are vaccinated (scenario 1): it takes 2 to 2.5 additional years to reach the same level of prevalence of shedders and 1.5 to 2 additional years to reach the same level of environmental load, although the two strategies only differ in the initial action of the control programme. Thus, from an epidemiological point of view, scenario 3, seems not the best strategy. In contrast, over the 10-year vaccination period of scenario 1, the mean prevalence of shedders and environmental bacterial load are decreased by 10 and 7 respectively. Although after 10 years of vaccination, the *C. burnetii* infection is still present in 58% of case herds, the vaccination of both heifers and cows from the start of the programme and for many years is in our study the most effective strategy. It has to be highlighted that the results of our study depend on the model structure and parameters values. The model represented the heterogeneity of shedding which is known to affect infection dynamics and hence the interventions efficacies in many diseases [104]. Indeed, model parameters governing

the shedding levels strongly influenced the *C. burnetii* dynamics (see Courcoul et al.[14]). Moreover, parameter values were inferred or calibrated from field data of naturally infected dairy cattle herds [35]. Thus, we took into account the latest knowledge on *C. burnetii* infections.

The probability of infection for an efficiently vaccinated susceptible cow p_v was quantified based on Guatteo et al. [61]. As the confidence interval of this parameter was wide, we studied the influence of this parameter value on the model outputs. Although the mean shedder prevalence was highly influenced by the value of p_v, the mean environmental bacterial load (which indirectly represents the infection risk for both animals and humans) decreased by roughly the same rate regardless of the parameter value. This is likely because the efficiently vaccinated animals shed in decreased quantities. Therefore, irrespective of whether the mean prevalence of vaccinated shedders remains high, the prevalence of high shedders was reduced, with a major impact on the environmental load. This result has also been described by Lu et al. [92] who showed that, to reduce the *Salmonella* prevalence in the long term, highly effective vaccines lowering the infectiousness would be a better choice than highly effective vaccines reducing susceptibility. Interestingly, whereas the environmental bacterial load was hardly sensitive to p_v (infection probability for efficiently vaccinated cows), the extinction rate was. Therefore, if the vaccine is to be used for eradication of *C. burnetii* from infected farms, both susceptibility and infectiousness have to be determined more accurately for the model to be used for prediction purposes or decision support. According to Rousset et al. [136], the lowest shedding level in vaginal swabs was shown to be more frequent in vaccinated than non vaccinated goats. However, further studies are needed to determine if a decrease of infectiousness is observed for all the vaccinated animals or only for the efficiently vaccinated ones and to quantify this decrease in all the shedding routes.

It should be noted that the extinction rate is highly influenced by the effect of vaccination on the susceptibility, the level of shedding and the mortality rate of the bacterium in the environment[3], which are all uncertain variables in our model. This extinction rate should then be interpreted with caution and used to compare different control strategies within the model. However, the behavior of the extinction rate suggests that it may be difficult and takes time to get free from *C. burnetii* within a herd.

[14] Courcoul A., Monod H., Nielen M., Klinkenberg D., Hogerwerf L., Beaudeau F., Vergu E., Modelling of the heterogeneity of shedding in the within herd *Coxiella burnetii* spread and identification of related key parameters through a sensitivity analysis, *submitted for publication* in Journal of Theoretical Biology.

In conclusion, our modeling approach showed that a long term yearly vaccination will reduce infection risk in vaccinated herds, but an additional cost-benefit analysis considering the economic aspects of control programmes is needed to design an optimal control strategy.

6. Acknowledgements

The authors would like to thank Alain Joly for fruitful discussions on the epidemiological model and vaccination strategies. The collaboration between the Faculty of Veterinary Medicine of Utrecht and the French National Institute for Agricultural Research (INRA) was financially supported by the Netherlands Organisation for Scientific Research (NWO) and the French Ministry of Foreign and European Affairs through the Van Gogh Programme.

7. Supplementary material

Tables 4.2 and 4.3 provide the definitions and values of all the parameters of the model.

Table 4.2. Description of the model parameters for the herd demography and their values used for simulations.

Parameters		Standard value
Replacement rate ($year^{-1}$)		0.355
Culling rate ($week^{-1}$)	lactation 1	0.0057
	lactation 2	0.0052
	lactation 3	0.0065
	lactation 4	0.0067
	lactations 5&6	0.0161
Probability distribution of the lactation numbers of the cows at the start of simulation	lactation 1	0.337
	lactation 2	0.252
	lactation 3	0.173
	lactation 4	0.11
	lactation 5	0.088
	lactation 6	0.04
Calving-calving interval (weeks)		55
Dry period (weeks)		8
Non gestation period (weeks)		15

Table 4.3. Definitions of the epidemiological model parameters and their values used for simulations.

Parameter		Definition	Value
m (week^{-1})		Transition rate I_1 => S and I_1V_e => SV_e	0.7 [a]
q (week^{-1})		Transition rate I_1 => (I_2 or I_3) and I_1V_e => (I_2V_e or I_3V_e)	0.02 [a]
p/p		Proportion of cows going from I_1 to (I_2 or I_3) and becoming I_3 and going from I_1V_e to (I_2V_e or I_3V_e) and becoming I_3V_e	0.5
r_1 (week^{-1})		Transition rate I_2 => C_1 and I_2V_e => C_1V_e	0.2 [a]
r_2 (week^{-1})		Transition rate I_3 => C_1 and I_3V_e => C_1V_e	0.02
s (week^{-1})		Transition rate C_1 => I_2 and C_1V_e => I_2V_e	0.15 [a]
τ (week^{-1})		Transition rate C_1 => C_2 and C_1V_e => C_2V_e	0.0096
μ (week^{-1})		Mortality rate of *C. burnetii*	0.2 [a]
probav		Probability of abortion after a transition S => I_1, C_1 => I_2 and C_2 => I_2	0.02
ρ^{mf}		Proportion of bacteria shed through mucus/faeces filling the environment compartment	0.2
ratio ρ^{milk}/ρ^{mf}		ρ^{milk} = proportion of bacteria shed through milk filling the environment compartment	0.125
α	milk	Probability distribution of the shedding routes for the I_1 cows	0.31[b]
	mucus/feces		0.62[b]
	milk+mucus/feces		0.07[b]
β	milk	Probability distribution of the shedding routes for the I_2 cows after 4 weeks post-calving	0.61[b]
	mucus/feces		0.33[b]
	milk+mucus/feces		0.06[b]
β_{calv}	milk	Probability distribution of the shedding routes for the I_2 cows in the 4 first weeks post-calving	0.14[b]
	mucus/feces		0.5[b]
	milk+mucus/feces		0.36[b]
γ	milk	Probability distribution of the shedding routes for the I_3 cows after 4 weeks post-calving	0.83[b]
	milk+mucus/feces		0.17[b]
γ_{calv}	milk	Probability distribution of the shedding routes for the I_3 cows in the 4 first weeks post-calving	0.25[b]
	milk+mucus/feces		0.75[b]
Q1	low level	Probability distribution of the shedding levels for all the I_1 and for the I_2 shedding in mucus/faeces after 4 weeks post-calving	0.85[b]
	mid level		0.15[b]
	high level		0[b]

$Q2$	low level	Probability distribution of the shedding levels for the I_2 shedding in milk after 4 weeks post-calving	0.4[b]
	mid level		0.5[b]
	high level		0.1[b]
$Q3$	low level	Probability distribution of the shedding levels for all the I_2 in the 4 first weeks post-calving	0.25[b]
	mid level		0.25[b]
	high level		0.5[b]
$Q4$	low level	Probability distribution of the shedding levels for the I_3 shedding in mucus/faeces after 4 weeks post-calving	0.6[b]
	mid level		0.4[b]
	high level		0[b]
$Q5$	low level	Probability distribution of the shedding levels for all the I_3 shedding in milk and for the I_3 shedding in mucus/faeces in the 4 first weeks post-calving	0.15[b]
	mid level		0.6[b]
	high level		0.25[b]
Qty (units of environment)	low level	Quantity of bacteria released by shedders in low, mid and high levels respectively	1/3000
	mid level		1/30
	high level		1
$Q1V_e$	low level	Probability distribution of the shedding levels for all the I_1V_e and for the I_2V_e shedding in mucus/faeces after 4 weeks post-calving	1
	mid level		0
	high level		0
$Q2V_e$	low level	Probability distribution of the shedding levels for the I_2V_e shedding in milk after 4 weeks post-calving	0.9
	mid level		0.1
	high level		0
$Q3V_e$	low level	Probability distribution of the shedding levels for the I_2V_e in the 4 first weeks post-calving	0.5
	mid level		0.5
	high level		0
$Q4V_e$	low level	Probability distribution of the shedding levels for all the I_3V_e shedding in mucus/faeces after 4 weeks post-calving	1
	mid level		0
	high level		0
$Q5V_e$	low level	Probability distribution of the shedding levels for all the I_3V_e shedding in milk and for the I_3V_e shedding in mucus/faeces in the 4 first weeks post-calving	0.75
	mid level		0.25
	high level		0
ratio p_v/p	standard value	Ratio between the transition rate $SV_e \Rightarrow I_1V_e$ and the transition rate $S \Rightarrow I_1$	0.21[c]
	bounds of the 95% CI tested for scenario 1		0.05 [c]
			0.9 [c]

[a]from Courcoul et al. Proc Biol Sci. (2010)

[b]calibrated to match field data (R. Guatteo 2009, personal communication)

[c]from Guatteo et al. Vaccine (2008)

CHAPTER 5

GENERAL DISCUSSION

Picture : M. Bédué

The aim of this thesis was to develop a model representing the spread of *C. burnetii* within a dairy herd in order to better understand the course of infection in cattle and to enlighten decision makers on the effectiveness of control measures, since this bacterium poses a problem for both human and animal health. This work was done in keeping with the EFSA recommendations which recently highlighted the need to objectively assess relevant epidemiological parameters and the effectiveness of control options for *C. burnetii* infection in domestic ruminants populations [39]. We focused on the within-herd *C. burnetii* spread. Although this scale may seem inappropriate for the study of the bacterial spillover from animal populations to humans, it is crucial to rigorously explore dynamics at finer scales before focusing on the whole dynamical process. More precisely, before representing the infection spread at a regional scale, understanding the within-herd infection dynamics is critical. The speed and trends of the within-herd infection spread, the heterogeneity related aspects between animals and/or farms are examples of key points that have to be checked before developing an inter-herds or an animal-human model of *C. burnetii* transmission.

Although Q fever is a European burning issue (mainly because of the current Dutch epidemic), the model we developed was to our knowledge the first epidemic model dealing with *C. burnetii* spread. Our study involved three main steps:

1. First, the model was conceptualised and the inference was performed based on field data in a Bayesian framework. The inference process was not only a prerequisite of the future use of the model, but it also allowed to quantify parameters having a biological meaning (e.g. probability of infection in a chronically infected herd, duration of shedding and non shedding periods, etc…) and then to better understand the natural course of the infection in a dairy cattle herd.

2. Then, after the model was made more complex and realistic by including variability within and between animals in the shedding duration, routes, and intensity, the factors most influencing the infection dynamics were determined through a sensitivity analysis. This step was necessary for linking the model uncertainty to some of the epidemic parameters: this allowed the model improvement by highlighting the need of an accurate quantification of uncertain but influential parameters. Moreover, the identification of influential mechanisms was part of the understanding of the infection process. It also plays a role in the definition of effective control strategies, by directing interventions towards the most vulnerable facets of the disease transmission.

3. At last, we tested by simulation the long term effectiveness of three different vaccination strategies in reducing the shedders prevalence, number of abortions, environmental

bacterial load, and in leading to extinction of infection. By taking into account the available knowledge on the vaccine effect, we thus determined the impact of vaccination according to the duration of the vaccination programme and the composition of the vaccinated population.

In section 1 of the general discussion the major findings of this thesis will be summarized. Then, section 2 will discuss the advantages and limits of the modelling approach which has been used, and section 3 the available and required data for model conceptualization, calibration and validation. Lastly, in section 4 a few implications and perspectives of this thesis work will be provided.

I- Major findings

The preliminary main achievement of this thesis was the elaboration of the first model in the literature for the study of *C. burnetii* spread within a dairy cattle herd and the effectiveness of different measures to control it.

First, the model constituted the basis for the exploration of heterogeneity related aspects. As highlighted by the available data, Q fever is characterised by a large heterogeneity both between herds and between animals. It was already observed that some infected herds were asymptomatic while others exhibited many abortions. Here, we showed that even for apparently similar herds (i.e. chronically infected herds without any obvious clinical sign attributable to Q fever), the infection dynamics was variable: intermittency of shedding was rare to usual according to the herd; few herds were characterized by a low probability of infection and then a slow bacterial spread, while others had a quite high probability of infection and then a faster infection dynamics. This heterogeneity in probability of infection was linked to the environmental bacterial load, which was variable between herds. In addition, within a herd, a high variability in the shedding routes, duration and levels of shedder cows was suggested by our data and also discussed in the literature. This heterogeneity of shedding was a key mechanism in the infection process: the most influential parameters were shown to be the probabilities governing the levels of shedding, especially for mucus/faeces shedders. Besides, seronegative infectious animals (I^-) and seropositive ones (I^+) were shown to have different patterns of shedding. I^- shed on average for a shorter period of time than I^+, more often in mucus only and almost exclusively at low titres. I^+ shed preferentially in milk only and quite often at moderate or high titres. The transition from one type of shedder to the other one was

rare (i.e. in chronically infected herds, the transition probability from *I-* to *I+* was shown to be very low). Due to this low estimated transition probability and to the high estimated transition probability from *I-* to susceptible animals (*S*), the simulated number of *I-* was higher than the number of *I+*, especially in the first years of infection. This could partly explain why the distribution of the levels of bacteria shed in mucus/faeces by the *I-* had a stronger impact on the model outputs than the one shed in mucus/faeces by the *I+*.

The parameters impacting the most the infection dynamics were also identified. Some physiological parameters related to the intermittency of shedding (i.e. transition probability from seropositive non-shedders to seropositive shedders) or to the transition from one type of shedder to another one (i.e. transition probability from seronegative shedders to seropositive ones) played a non-negligible role. However, the most influential parameters were associated to the probabilities governing the levels of shedding, especially for mucus/faeces shedders, as already mentioned above, and to the characteristics of the bacterium in the environment (i.e. proportion of bacteria shed through mucus/faeces reaching the environment and mortality rate of *C. burnetii*). Interventions impacting those key parameters would be of great interest. Therefore, control measures leading to (i) a decrease in the quantities of bacteria shed, especially in vaginal mucus and/or faeces, (ii) a decrease in the probability of shedding again for an infected non-shedder animal or to (iii) a decrease in the life expectancy of *C. burnetii*, could be a priori effective control strategies. As vaccination and environmental measures (e.g. increased cleaning and disinfection) are susceptible to respectively decrease the quantities of bacteria shed and the life expectancy of the bacterium, they could be promising interventions.

The relative effectiveness of three vaccination strategies was determined in infected dairy herds subject to at least three abortions in the previous year. Vaccinating cows and heifers for three years only was ineffective in the long run. The probability of extinction of the infection was low using this scheme. Thus, although the prevalence of shedders, the environmental bacterial load, and the number of abortions decreased during the vaccination programme, they increased again after the campaign was ceased. In contrast, a 10-year vaccination period for both cows and heifers allowed to considerably decrease the mean prevalence of shedders, environmental bacterial load, and number of abortions and even to eradicate the infection in a non-negligible number of cases. However, when only heifers (instead of cows and heifers) were vaccinated at the beginning of this 10-year vaccination programme, the decrease was much slower and it took about two additional years to reach the same level of shedders prevalence or environmental load.

II- Comments on the modelling approach, inference and model analysis

1. Choice of the mathematical formalism

To represent the spread of *C. burnetii* within a dairy cattle herd, we developed a stochastic individual-based model in discrete time with a time step of one week. As the mean size of the host population was low (around 50 cows), it was appropriate to consider a stochastic approach. For each individual, all the transitions between health states as well as the determination of the shedding routes and quantities of bacteria shed in the environment were supposed stochastic. The individual-based scale was preferred in order to allow the estimation of parameters of transitions between health states in the presence of bidirectional transitions. Moreover, this representation enabled to consider the lactation and gestation cycle of cows, which interfered with the epidemic process. As an example, the probability distributions of the levels of shedding were variable with respect to the moment of calving. Therefore, an individual-based model was an easy way to take into account interactions between demography and epidemiology by recording the individual life history of each cow. A one-week time step was chosen since no transition was assumed to occur in less than seven days. However, this choice was driven by the data set configuration (in data set A, samplings are performed every week) and it could be interesting to check this assumption by sampling cows, especially intermittent shedders, every two or three days.

2. Choice of the model structure

The complete model we developed (i.e. the variant including heterogeneity related aspects) had a SIR-like structure with three different kinds of I (I-, I+ and $I^{+milk\ pers}$). This model was characterised by transitions in both directions between S and I- and between I+ and R (R health state was called C+ in the model including the heterogeneity of shedding and C_1 in the model including vaccination). Several issues were encountered. First, the S state comprised real susceptible animals but also apparently susceptible animals: seronegative non-shedders which were already infected became I- and went back to S. Although they were seronegative, the latter could have developed a cell-mediated immunity response to *C. burnetii* which means that they were not naïve to the infection anymore. The transition probability from S to I- is then a mix between an infection probability and a re-infection probability. However, using the current

diagnostic tests it was not possible to differentiate primary infected from re-infected animals. Cell-mediated immunity tests (i.e. skin tests) are under development and would have been of great interest to make this distinction.

We represented in the model two types of infectious animals, *I-* and *I+* (considering *I+* and $I^{+}_{milk\ pers}$ all together) with two different shedding patterns. However, individual factors such as age, genetics, immunity, or other management factors that might predispose a cow to fall into one of these two categories are still unknown. When developing our model, we considered the importance of the humoral immunity response, which was assumed to be the main differential factor leading to the distinction between *I-* and *I+*. An additional difference between these two infectious states was that when in the first one, an animal had the possibility to clear the infection, while when in the second he stayed infected and alternated shedding and non shedding episodes. When representing *C. burnetii* spread in vaccinated herds, we had to define in the model the type(s) of *I* for vaccinated and therefore seropositive animals. Two modeling options were considered: (i) to still assume that the humoral immunity response was the differential factor between the two types of *I*: hence, all vaccinated animals have the *I+* shedding pattern, or (ii) to assume that the humoral immunity response was not the only factor enabling the distinction between the two types of *I*: two different shedding patterns could exist in vaccinated animals too[15]. In order to choose the "best" option, we looked for field data on shedding routes distributions in vaccinated cows. In Guatteo et al [61], the shedding route distribution of 18 cows, susceptible when vaccinated and monitored for the following one-year period, was not significantly different from the one of *I-* or from the one of *I+* in non-vaccinated herds. Thus, it was impossible to determine if the type of shedder could be determined according to the humoral response, the time from infection or another factor. Between the two modeling options, we finally chose the second one: although all vaccinated animals were assumed seropositive, two types of *I* were assumed to exist for vaccinated animals and represented in the model (i.e. I_1 and $I_{2/3}$). Further studies are highly needed to define (i) individual factors that determine the shedding pattern and then the type of shedder (cell-mediated immunity tests could be useful for this part) and (ii) the impact of vaccination on the shedding patterns.

Another option would be to define different types of shedders regardless of the presence or absence of antibodies. We could also assume that infected individuals which succeed in clearing the infection become resistant and do not get infected again. A new conceptual model could be

[15] we assumed in this case that one type of *I*, I_1 has a shedding pattern similar to *I-*, and that the other one, I_2, is similar to *I+*.

proposed (Figure 5.1): this new version would have the advantage of distinguishing real susceptible (*S*) from apparently susceptible (*R*) individuals. It would also match the opinion of some experts who consider that humoral immunity response is not a key point of the infection process. However, according to the data set A, we had no other choice when making inference than gathering real and apparently susceptible individuals into a unique category and defining classes of shedders based on their serological status.

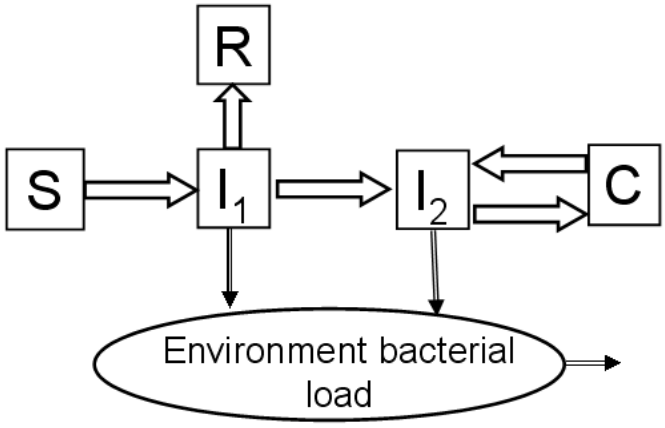

Figure 5.1. Flow diagram representing a possible description of the spread of *C. burnetii* within a cattle herd. The health states are: *S*, the real susceptible individuals, I_1, the shedder cows which are able to clear the infection, *R*, the resistant animals, I_2, the chronically infected cows which are shedding, and *C*, the chronically infected cows which are not shedding.

As a last element of the model structure, the route of *C. burnetii* transmission represented in the model was the inhalation of bacteria from the environment. Based on current knowledge, this infection route is the main one. However, if in the future, ticks, wildlife, or other transmission pathways are shown to have a non-negligible role in the infection dynamics, the model structure will need to be adapted.

3. Estimation of main epidemiological parameters

We used a Bayesian framework to estimate from field data the probabilities of transitions between health states as well as the parameters linked to the shedding and survival of *C. burnetii* in the environment. This approach gave us the possibility (i) to deal with missing data, (ii) to account for both previous knowledge about *C. burnetii* (mainly concerning the life expectancy of the bacteria in the environment and the proportions of different health states within an infected herd), and (iii) to take into account between-herd heterogeneity by considering some transition parameters as herd dependent. The results of the estimation work were satisfactory: for the large majority of parameters, a good convergence of the Monte

Carlo Markov chains was achieved, the posterior distributions obtained were biologically plausible, and the model was able to properly reproduce the observed data.

It would have been possible to take into account the uncertainty of observations in a different way. In the Bayesian network we developed, the observed health state of cow *i* at time *t*, $O_t^{(i)}$, is a random qualitative nominal variable which can take the values *S*, *I-*, *I+* or *R*, according to the real health state of this cow at the same time, $R_t^{(i)}$, and probabilities given by the sensitivities and specificities of the diagnostic tests. More technically, this can be viewed as a hidden Markov model, where the modelled system is assumed to be Markovian with hidden (unobserved) states, here variables $R_t^{(i)}$. However, as discussed in section III.6 of chapter 2, a possible way to more accurately account for the uncertainty of the observed health states is to consider that this uncertainty would differ for each observation as a function of the quantitative results provided by the diagnostic tests (Optical Densities for the ELISA and Ct values for the PCR). Indeed, this would potentially increase the accuracy since probabilities linking the hidden and observed layers would no more depend on sensitivities and specificities, which are already averaged values over the whole population, but rather directly on individual information. We could assume that the uncertainty on an observation is greater when the quantitative test result is close to the positivity threshold than when it is far away. To investigate this avenue, probability distributions of OD and Ct conditionally to the real health state of individuals were modelled (Figure 5.2) and parameterized according to the ranges of observed OD and Ct values and to the sensitivities and specificities of the diagnostic tests. As an example, to determine the probability that a cow has a given value of OD knowing that it has antibodies (green line of Figure 5.2.a), we first determined the range of observed OD values [-70 to 630]. The ELISA had an assumed sensitivity of 85% and the positivity threshold was 40. We assumed that the mode for the OD value for animals with antibodies is 190. The probability distribution had therefore to fulfil two conditions: its mode had to be 190 and the area under the curve for OD values above 40 had to represent 85% of the total area under the curve. In this way, we determined the probability distribution parameters. During our thesis work, we did not have enough time to further investigate this way of taking into account the uncertainty on observations but it could represent an interesting extension.

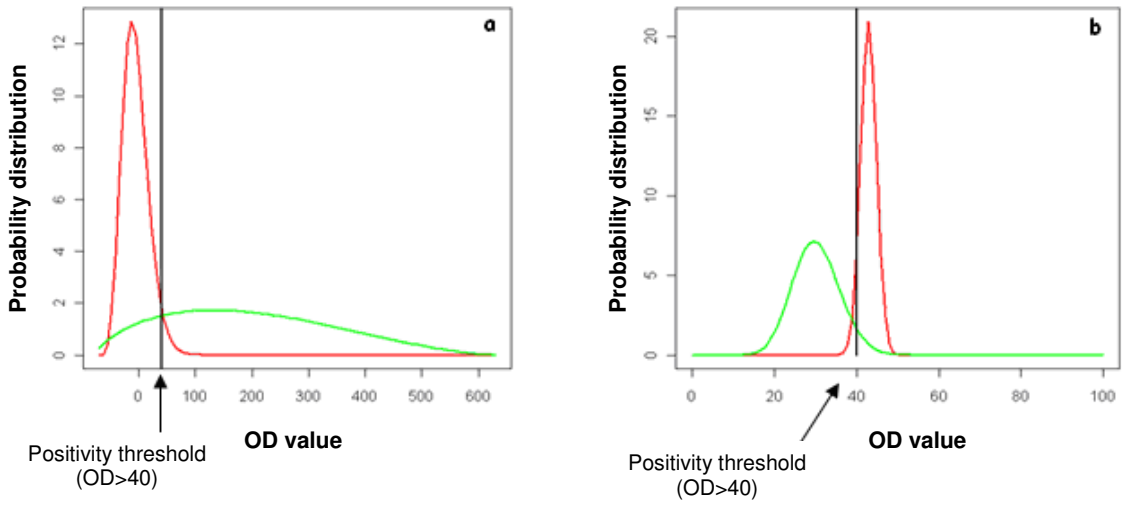

Figure 5.2 - a. Probability that a cow has a given value of Optical Density knowing that (i) it does not have any antibodies (red line), (ii) it has antibodies (green line);
- b. Probability that a cow has a given value of Cycle threshold knowing that (i) it is not a shedder (red line), (ii) it is a shedder (green line)

We could also wonder if all the epidemiological parameters of the complete model (i.e. model including the shedding routes and levels developed in chapter 3) could have been estimated from data set A in this Bayesian framework. In this case, more parameters would have to be estimated and it is not sure that the convergence of the MCMC would be achieved, although information about shedding routes and levels from data set A could be used. A compromise would be perhaps to perform a sensitivity analysis on the complete model first, to fix its non-influential parameters at their most plausible values (from the literature or based on expert opinion), and then to only estimate model parameters influencing the infection dynamics.

4. Sensitivity analysis

To identify the parameters that mostly contributed to the model output variability, we performed a sensitivity analysis considering 19 epidemiological parameters of the complete model. We defined four levels per factor directly related to the heterogeneity of shedding and two levels per other factors, and used fractional or complete fractional designs. As our model was stochastic, we studied the variability of the mean and standard deviation of 30 model repetitions for seven outputs of interest. This allowed to differentiate the variability due to the inherent model stochasticity from the variability due to the variability in input factors, and to focus only on the latter. Besides, the seven outputs of interest were dynamical and recorded every week for a period of five years. Therefore, we first performed a PCA to summarize the behaviour of the outputs over the whole simulated period and then an ANOVA to compute

sensitivity indices. This approach developed by Lamboni et al. [84] was preferred to Sobol's method and FAST which are also variance-based methods allowing the calculation of sensitivity indices. In these methods, the range of variation of each factor has to be defined by a continuous probability distribution. In our study, as some factors were not scalars but probability distributions, it seemed difficult to describe their range of variation in a continuous way. Besides, for Sobol's method and FAST, model outputs have to be punctual and not dynamical. Thus, the PCA followed by ANOVA method we used seemed better suited to our needs.

5. Simulation of control strategies

In our work, we tested the relative effectiveness of three vaccination strategies in herds characterized by at least three abortions in the last year. The modelling approach had the major advantage that every parameter of the infection could be monitored over time. As an example, the environmental bacterial load was recorded weekly during the control programme, whereas this factor is not easily available in field studies. In addition, the model developed in this thesis work is a flexible tool that can be easily adapted to explore other research questions related to *C. burnetii* spread and control. It would be possible to test additional control strategies (e.g. environmental measures, specific culling, etc…) as well as the same interventions but for different initial conditions (e.g. vaccination in non infected herds, in herds after the 1st abortion attributable to Q fever occurred, etc…).

III- Available and required data for model conceptualization, inference and validation

A key aspect of our modelling work was the model elaboration and the estimation of its parameters from field data. Data set A was of great value as 235 cows from five chronically infected herds were sampled from one to five times over a one-month period. At each sampling time, the serological status of the cow was determined as well as its shedding pattern (i.e. shedding routes and levels) if the cow was shedding. This information allowed us to estimate the parameters of transition between health states and to calibrate the representation of the heterogeneity of shedding when incorporating the shedding routes and levels in the model. However, this data came from infected herds without any obvious clinical sign attributable to Q fever and was used to also describe the infection spread in herds assumed to experience abortions due to Q fever. It would therefore be useful to collect data from herds exhibiting

clinical signs in order to check whether their disease dynamics would be quite similar to the one of herds without any visible sign.

Concerning the initial point of the infection, no data was available. It is likely that the parameters of transitions between health states at the very beginning of the pathogen spread differ from those of infected herds where the bacterium has been present for a long time. As an example, in this latter type of herds (i.e. chronically infected herds), the estimated value of the transition probability from *I-* to *I+* was almost nil, whereas they were many *I+* cows. Hence, we could imagine that, in the early stages of infection, some cows get infected, become *I-* and then *I+*, during a relatively short time span. The probability of transition from *I-* to *I+* would therefore be probably higher in recently infected herds compared to chronically infected herds, where almost no transitions between these two states are observed. In the field, detection of the infection occurs late, because it is based on the occurrence of clinical signs. Experimental studies are therefore the only way to monitor early stages of *C. burnetii* infections. Such studies are logistically complex and costly as *C. burnetii* has to be handled in P3 laboratories. However, if conceived, they would allow determining the speed of the infection spread at the beginning of the process and also to clarify the model parameterization. Specific factors that could influence the individual response to *C. burnetii* and pattern of infection could also be studied. As an example, the gestation status is an influential factor: pregnant goats experimentally challenged with *C. burnetii* are often seronegative and do not shed until they abort or kid (H.J. Roest, personal communication). This kind of mechanism has not been included in the model yet and further studies would therefore be highly needed to specify the first stages of infection.

As previously mentioned, there would be a need of data about cell-mediated immunity: this could help distinguishing the real from the apparently susceptible animals and the different types of *I* individuals, and thus to update the model structure. This type of immunity was indeed reported to play a role in the clinical expression of *C. burnetii* infection [73]. Moreover, skin tests were recently used to determine the interest of an annual booster in vaccinated cattle by assessing the level of different immune markers [133]. However, to our knowledge, no study has investigated yet the possible link between cell-mediated immunity and shedding pattern.

Another beneficial area of research would be the quantification of viable *Coxiella* in the environment. It could help quantifying more accurately the model parameters directly related to the environmental bacterial load (proportion of bacteria shed through mucus/faeces reaching the environment and mortality rate of *C. burnetii*) and to check that the mathematical

expression giving the probability of infection with respect to the environmental bacterial load is appropriate. As previously explained (see section IV.1. of the introduction), current methods monitoring the environmental bacterial load do not allow easily quantifying the risk of infection. PCR on dust samples can be performed, but it is difficult to correlate the result of this test with the environmental load of bacteria actually contributing to the infection. Indeed, pathogens can be trapped in dust depositions, which decreases their availability and hence the probability of airborne infection of cattle. Therefore, both the pathogen viability and the probability of inhalation of the contaminated dust should be determined. A promising technique is the PCR on air samples. The probability that *C. burnetii* could be inhaled has not to be quantified as bacteria are already in suspension in the air. A protocol of air sampling in infected dairy goat farms is currently led at the University of Utrecht (the Netherlands) jointly by the Faculty of Veterinary Medicine and the Institute for Risk Assessment Sciences.

This work also highlighted a lack of knowledge on the effect of control strategies. Regarding the vaccination, the relative risk of infection for animals vaccinated when susceptible and non-pregnant was quantified [61]. Although its confidence interval was wide (i.e. 95%CI of 0.05-0.90), this estimation guided the choice of numerical values for one of our model parameters. However, the consequences of vaccination on shedding routes and levels have not been evaluated yet. Also, no information is currently available on the consequences of environmental measures on the viability of *C. burnetii* in the farm. These knowledge gaps would need to be filled in before optimal control strategies could be defined.

Model validation is a key step before using model predictions to guide public health decision makers. Sensitivity analysis, model assumptions relevancy, and checking of the concordance between model conclusions and expert opinions are part of the process. The main step consists in confronting model outputs to independent data sets [163]. In our case, it was impossible to perform this confrontation, especially for the environmental bacterial load since no additional data was available. Nevertheless, we ran 100 repetitions of the complete model (i.e. including shedding routes and levels) for a period of 5 years and compared the mean simulated seroprevalence and prevalence of shedders for the last point of the time series with published data. The concordance was satisfactory: the simulated seroprevalence was equal to 35% on average (23.3% and 47.8% for the 25^{th} and 75^{th} percentiles) compared to 40% on average (25% and 51% for the 25^{th} and 75^{th} percentiles) in 56 naturally infected French herds with abortions due to Q fever [149]. The mean simulated prevalence of shedders was 35.5% (0 and 61.7% for the percentiles 2.5^{th} and 97.5^{th}) whereas in the field, the prevalence of shedders was 45.5% for 242 cows from 31 naturally infected French herds with abortions due to Q fever [57] and

38.9% in the 95 cows from three French asymptomatic dairy herds [131]. However, this data is neither repeated over time, nor does it allow following the infection spread. It would have been possible to split the data set A into two subsets (e.g. data from three herds on the one hand and from the two other herds on the other hand). One subset would have been used to estimate model parameters and the other to validate the model. Though, due to missing data, information available in data set A was limited. Besides, as some estimated parameters were assumed to be herd-dependent, the whole data set A had to be used for parameter estimation.

Currently, an 18-month follow-up of 100 naturally infected dairy cattle herds with abortions is led in the west of France by the unit of Oniris-INRA "Bioaggression, Epidemiology and Risk Analysis in Animal Health" (Nantes, France). The aim of this study is to assess the impact of control strategies combining vaccination and/or antibiotherapy in field conditions. Criteria of effectiveness include results of real-time PCR in bulk tank milk performed every three months. One of four potential control strategies is implemented in each herd: (i) vaccination of heifers only, (ii) vaccination of heifers and cows, (iii) antibiotherapy before calving and drying up, (iv) vaccination of heifers and cows as well as antibiotherapy before calving and drying up. Interventions (i) and (ii) correspond to two vaccination strategies that we simulated with our model. It would then be possible to confront model outputs to these field data. Nevertheless, as the correlation between the titre in bacteria in bulk milk samples and the prevalence of milk shedders remains weak [58], the comparison between observed and simulated results seems difficult.

IV- Implications and perspectives

The work that was carried out during my PhD has two types of immediate consequences. Firstly, it highlights and helps to prioritize needs of research. As previously discussed, further studies aiming at assessing the environmental load of viable bacteria, at describing the first stages of the infection process, at determining the possible role of cell-mediated immunity in the evolution of individual health states, and at quantifying the vaccine effect on the shedding pattern would be of great value. Besides, identifying individual or environmental factors that lead to super-shedding events, especially in vaginal mucus or faeces, would be a key milestone in the understanding and control of the infection spread. Secondly, our work can guide farmers and decision makers in the choice and design of control programmes for Q fever in cattle. Vaccination is an effective way to decrease both the shedder prevalence and the environmental bacterial load (under the assumption that effectively vaccinated animals shed in decreased quantities). According to epidemiological considerations, both heifers and cows should be

vaccinated, as administrating vaccines to heifers only is less effective: on average, to reach a given level of shedder prevalence takes two additional years when vaccinating heifers only. Performing a comparative analysis of cost-effectiveness between these two types of vaccination strategies would be valuable for decision making. Even in presence of vaccination, eradication of infection seems difficult and most of the time, takes several years. Therefore, before stopping a vaccination programme, it is relevant to check for the absence of *C. burnetii* within the herd. If some shedders are still present, it is likely that the infection will spread again after the programme is stopped. PCR on bulk tank milk samples can be a way to perform this verification. As shedding is intermittent and as its routes are not concomitant, this diagnostic test would better be repeated over time. Other control measures, and especially environmental ones, such as increased farm cleaning and disinfection, seem promising. Decreasing the life time of *C. burnetii* in the environment as well as the proportion of bacteria reaching the environment (e.g. by rapidly destroying parturition or abortion products, by increasing housing cleaning around calving, etc…) could strongly impact the environmental load. The model we developed is an adaptable tool that would allow assessing the effectiveness of a broad range of other control strategies in different initial situations. In the next years, it is then crucial to regularly update this tool as new knowledge is produced.

Since small ruminants are often responsible for human infections, it seems relevant to adapt our model to represent *C. burnetii* spread within sheep and goat flocks. The demography of small ruminant populations differs from the one of cattle herds: kidding is most of the time synchronised. Besides, flocks can be very large (several hundreds and even thousands of animals per flock, as it is currently the case in the Netherlands). It would be interesting to determine if the flock size and type of herd management have an influence on the infection dynamics and to assess the effectiveness of control measures in this context. A collaborative work with the University of Utrecht in the framework of Lenny Hogerwerf's PhD should answer these questions soon.

As previously highlighted, Q fever is characterised by a large between-herd heterogeneity with some asymptomatic infected herds, while others exhibit many abortions. This variability in the occurrence and intensity of the clinical signs could be partly due to variability in *C. burnetii* strains involved in ruminants infections: in Arricau-Bouvery et al. [14], 36 different genotypes were identified among the 42 isolates from livestock and ticks investigated. However, neither the virulence of the different *C. burnetii* strains nor possible interactions between them have been investigated yet. If it would be shown that *C. burnetii* strains have different virulence or

shedding characteristics or interact through cross-immunity, our model should be modified to account for multi-strain dynamics.

Lastly, a next step in the understanding and control of *C. burnetii* spread would be the study of the infection spread at the regional level. Likely, both cattle herds and small ruminant flocks on a given area have to be represented in order to understand the global infection dynamics and to evaluate the respective influence of herds and flocks on the infection risk. The role of purchase of animals, wind, and neighbouring contacts on the spread of *C. burnetii* between animal populations has also to be determined. At last, the effectiveness of control measures at a regional or national scale should be assessed, which would provide evidence for decision making.

REFERENCES

[1] Acedo L., Diez-Domingo J., Morano J.A., Villanueva R.J., Mathematical modelling of respiratory syncytial virus (RSV): vaccination strategies and budget applications, Epidemiol Infect. (2010) 138:853-860.

[2] AFSSA, Fièvre Q : Rapport sur l'évaluation des risques pour la santé publique et des outils de gestion des risques en élevage de ruminants. 2003. [95 pp.]. Available online: http://www.afssa.fr/Documents/SANT-Ra-fievreQ.pdf

[3] Agger J.F., Christoffersen A.B., Rattenborg E., Nielsen J., Agerholm J.S., Prevalence of *Coxiella burnetii* antibodies in Danish dairy herds, Acta Vet Scand. (2010) 52:5.

[4] Agur Z., Cojocaru L., Mazor G., Anderson R.M., Danon Y.L., Pulse mass measles vaccination across age cohorts, Proc Natl Acad Sci U S A. (1993) 90:11698-11702.

[5] Aitken I.D., Clinical aspects and prevention of Q fever in animals, European J. Epidemiology. (1989) 5:420-424.

[6] Alba A., Sanchez-Cabre D., Badiella L., Allepuz A., Napp S., Garcia I., Casal J., Evolution of the BSE epidemic in Catalonia (1990-2015) based on a stochastic model, Vet J. (2010) 184:182-186.

[7] Andraud M., Rose N., Grasland B., Pierre J.S., Jestin A., Madec F., Influence of husbandry and control measures on porcine circovirus type 2 (PCV-2) dynamics within a farrow-to-finish pig farm: a modelling approach, Prev Vet Med. (2009) 92:38-51.

[8] Angelakis E., Raoult D., Q fever, Vet Microbiol. (2010) 140 (3-4):297-309.

[9] Arricau-Bouvery N., Souriau A., Moutoussamy A., Ladenise K., Rodolakis A., Etude de l'excrétion de *Coxiella burnetii* dans un modèle expérimental caprin et décontamination des lisiers par la cyanamide calcique, Renc. Rech. Ruminants. (2001) 8:153-156.

[10] Arricau-Bouvery N., Souriau A., Lechopier P., Rodolakis A., Excretion of *Coxiella burnetii* during an experimental infection of pregnant goats with an abortive goat strain CbC1, Ann N Y Acad Sci. (2003) 990:524-526.

[11] Arricau-Bouvery N., Souriau A., Lechopier P., Rodolakis A., Experimental *Coxiella burnetii* infection in pregnant goats: excretion routes, Vet Res. (2003) 34:423-433.

[12] Arricau-Bouvery N., Rodolakis A., Is Q fever an emerging or re-emerging zoonosis?, Vet Res. (2005) 36:327-349.

[13] Arricau-Bouvery N., Souriau A., Bodier C., Dufour P., Rousset E., Rodolakis A., Effect of vaccination with phase I and phase II *Coxiella burnetii* vaccines in pregnant goats, Vaccine. (2005) 23:4392-4402.

[14] Arricau-Bouvery N., Hauck Y., Bejaoui A., Frangoulidis D., Bodier C.C., Souriau A., Meyer H., Neubauer H., Rodolakis A., Vergnaud G., Molecular characterization of *Coxiella burnetii* isolates by infrequent restriction site-PCR and MLVA typing, BMC Microbiol. (2006) 6:38.

[15] Arricau Bouvery N., Souriau A., Lechopier P., Rodolakis A., Experimental *Coxiella burnetii* infection in pregnant goats: excretion routes, Vet Res. (2003) 34:423-433.

[16] Astobiza I., Barandika J.F., Hurtado A., Juste R.A., Garcia-Perez A.L., Kinetics of *Coxiella burnetii* excretion in a commercial dairy sheep flock after treatment with oxytetracycline, Vet J. (2009) 184:172-175.

[17] Backer J.A., Hagenaars T.J., van Roermund H.J., de Jong M.C., Modelling the effectiveness and risks of vaccination strategies to control classical swine fever epidemics, J R Soc Interface. (2009) 6:849-861.

[18] Bailey R.A., Design of comparative experiments, Cambridge University Press, Cambridge, 2008.

[19] Barlow J., Rauch B., Welcome F., Kim S.G., Dubovi E., Schukken Y., Association between *Coxiella burnetii* shedding in milk and subclinical mastitis in dairy cattle, Vet Res. (2008) 39:23.

[20] Berri M., Souriau A., Crosby M., Crochet D., Lechopier P., Rodolakis A., Relationships between the shedding of *Coxiella burnetii*, clinical signs and serological responses of 34 sheep, Vet Rec. (2001) 148:502-505.

[21] Berri M., Rousset E., Champion J.L., Arricau-Bouvery N., Russo P., Pepin M., Rodolakis A., Ovine manure used a a garden fertiliser as a suspected source of human Q fever, Vet Rec. (2003) 153:269-270.

[22] Berri M., Crochet D., Santiago S., Rodolakis A., Spread of *Coxiella burnetii* infection in a flock of sheep after an episode of Q fever, Vet Rec. (2005) 157:737-740.

[23] Berri M., Rousset E., Champion J.L., Russo P., Rodolakis A., Goats may experience reproductive failures and shed *Coxiella burnetii* at two successive parturitions after a Q fever infection, Res Vet Sci. (2007) 83:47-52.

[24] Berri M., Rekiki A., Boumedine K.S., Rodolakis A., Simultaneous differential detection of Chlamydophila abortus, Chlamydophila pecorum and *Coxiella burnetii* from aborted ruminant's clinical samples using multiplex PCR, BMC Microbiol. (2009) 9:130.

[25] Biberstein E.L., Riemann H.P., Franti C.E., Behymer D.E., Ruppanner R., Bushnell R., Crenshaw G., Vaccination of dairy cattle against Q fever (*Coxiella burnetii*): results of field trials, Am J Vet Res. (1977) 38:189-193.

[26] Bildfell R.J., Thomson G.W., Haines D.M., McEwen B.J., Smart N., *Coxiella burnetii* infection is associated with placentitis in cases of bovine abortion, J Vet Diagn Invest. (2000) 12:419-425.

[27] Bouvier A., Kobilinsky A., Monod H., PLANOR : an R library for the automatic generation of regular fractional factorial designs. 2010.

[28] Bruijnis M.R., Hogeveen H., Stassen E.N., Assessing economic consequences of foot disorders in dairy cattle using a dynamic stochastic simulation model, J Dairy Sci. (2010) 93:2419-2432.

[29] Buhariwalla F., Cann B., Marrie T.J., A dog-related outbreak of Q fever, Clin Infect Dis. (1996) 23:753-755.

[30] Cariboni J., Gatelli D., Liska R., Saltelli A., The role of sensitivity analysis in ecological modelling, Ecological Modelling. (2007) 203:167-182.

[31] Cauchemez S., Carrat F., Viboud C., Valleron A.J., Boelle P.Y., A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data, Stat Med. (2004) 23:3469-3487.

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

[33] Centers for Disease Control and Prevention [on line] http://wwwn.cdc.gov/travel/content/id/1769.aspx [consulted Accessed on 2009/12/21].

[34] Courcoul A., Ezanno P., Modelling the spread of Bovine Viral Diarrhoea Virus (BVDV) in a managed metapopulation of cattle herds, Vet Microbiol. (2010) 142:119-128.

[35] Courcoul A., Vergu E., Denis J.B., Beaudeau F., Spread of Q fever within dairy cattle herds: key parameters inferred using a Bayesian approach, Proc Biol Sci. (2010) 277:2857-2865.

[36] Diekmann O., Heesterbeek H., Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation, John Wiley & Sons, Chichester, UK, 2000.

[37] Durand M.P., Lacteal and placental excretion of *Coxiella burnetii*, agent of Q fever, in the cow. Importance and prevention, Bull Acad Natl Med. (1993) 177:935-945; discussion 945-936.

[38] ECDC, Risk assessment on Q fever. 2010. [40 pp.]. Available on line: http://ecdc.europa.eu/en/publications/Publications/1005_TER_Risk_Assessment_Qfever.pdf

[39] EFSA Panel on Animal Health and Welfare (AHAW), Scientific Opinion on Q Fever. EFSA Journal 2010; 8(5):1595, [114 pp.]. doi:10.2903/j.efsa.2010.1595. Available online: www.efsa.europe.eu

[40] England T., Kelly L., Jones R.D., MacMillan A., Wooldridge M., A simulation model of brucellosis spread in British cattle under several testing regimes, Prev Vet Med. (2004) 63:63-73.

[41] Enoe C., Georgiadis M.P., Johnson W.O., Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown, Prev Vet Med. (2000) 45:61-81.

[42] Enserink M., Infectious diseases. Questions abound in Q-fever explosion in the Netherlands, Science. (2010) 327:266-267.

[43] Ezanno P., Fourichon C., Viet A.F., Seegers H., Sensitivity analysis to identify key-parameters in modelling the spread of bovine viral diarrhoea virus in a dairy herd, Prev Vet Med. (2007) 80:49-64.

[44] Ezanno P., Fourichon C., Seegers H., Influence of herd structure and type of virus introduction on the spread of bovine viral diarrhoea virus (BVDV) within a dairy herd, Vet Res. (2008) 39:39.

[45] Ezanno P., Vergu E., Langlais M., Gilot-Fromont E., Modelling the dynamics of host-parasite interactions: basic principles, in: Springer (Ed.), New frontiers of molecular epidemiology of infectious diseases, in press.

[46] Ferguson N.M., Donnelly C.A., Woolhouse M.E., Anderson R.M., The epidemiology of BSE in cattle herds in Great Britain. II. Model construction and analysis of transmission dynamics, Philos Trans R Soc Lond B Biol Sci. (1997) 352:803-838.

[47] Fournier P.E., Marrie T.J., Raoult D., Diagnosis of Q fever, J Clin Microbiol. (1998) 36:1823-1834.

[48] Galvani A.P., May R.M., Epidemiology: dimensions of superspreading, Nature. (2005) 438:293-295.

[49] Garner M.G., Cowled B., East I.J., Moloney B.J., Kung N.Y., Evaluating the effectiveness of early vaccination in the control and eradication of equine influenza-A modelling approach, Prev Vet Med. (2010) in press.

[50] Garske T., Rhodes C.J., The effect of superspreading on epidemic outbreak size distributions, J Theor Biol. (2008) 253:228-237.

[51] Gelman A., Rubin D.B., Inference for iterative simulation using multiple sequences, Statist. Sci. (1992) 7:457-511.

[52] Gikas A., Kokkini S., Tsioutis C., Q fever: clinical manifestations and treatment, Expert Rev Anti Infect Ther. (2010) 8:529-539.

[53] Gilks W.R., Richardson S., Spiegelhalter D.J., Introducing Markov chain Monte Carlo, in: Chapman&Hall (Ed.), Markov chain Monte Carlo in practice, London, 1996, pp. 1-19.

[54] Gilsdorf A., Kroh C., Grimm S., Jensen E., Wagner-Wiening C., Alpers K., Large Q fever outbreak due to sheep farming near residential areas, Germany, 2005, Epidemiol Infect. (2008) 136:1084-1087.

[55] Glazunova O., Roux V., Freylikman O., Sekeyova Z., Fournous G., Tyczka J., Tokarevich N., Kovacava E., Marrie T.J., Raoult D., *Coxiella burnetii* genotyping, Emerg Infect Dis. (2005) 11:1211-1217.

[56] Greenslade E., Beasley R., Jennings L., Woodward A., Weinstein P., Has *Coxiella burnetii* (Q fever) been introduced into New Zealand?, Emerg Infect Dis. (2003) 9:138-140.

[57] Guatteo R., Beaudeau F., Berri M., Rodolakis A., Joly A., Seegers H., Shedding routes of *Coxiella burnetii* in dairy cows: implications for detection and control, Vet Res. (2006) 37:827-833.

[58] Guatteo R., Beaudeau F., Joly A., Seegers H., Assessing the within-herd prevalence of *Coxiella burnetii* milk-shedder cows using a real-time PCR applied to bulk tank milk, Zoonoses Public Health. (2007) 54:191-194.

[59] Guatteo R., Beaudeau F., Joly A., Seegers H., *Coxiella burnetii* shedding by dairy cows, Vet Res. (2007) 38:849-860.

[60] Guatteo R., Beaudeau F., Joly A., Seegers H., Performances of an ELISA applied to serum and milk for the detection of antibodies to *Coxiella burnetii* in dairy cattle, Revue De Medecine Veterinaire. (2007) 158:250-252.

[61] Guatteo R., Seegers H., Joly A., Beaudeau F., Prevention of *Coxiella burnetii* shedding in infected dairy herds using a phase I C. *burnetii* inactivated vaccine, Vaccine. (2008) 26:4320-4328.

[62] Hilbink F., Penrose M., Kovacova E., Kazar J., Q fever is absent from New Zealand, Int J Epidemiol. (1993) 22:945-949.

[63] Ho T., Htwe K.K., Yamasaki N., Zhang G.Q., Ogawa M., Yamaguchi T., Fukushi H., Hirai K., Isolation of *Coxiella burnetii* from dairy cattle and ticks, and some characteristics of the isolates in Japan, Microbiol Immunol. (1995) 39:663-671.

[64] Hogerwerf L., Van den Brom R., Roest H.J., Bouma A., Vellema P., Pieterse M., Dercksen D., Nielen M., Vaccination of dairy goat herds reduces *Coxiella burnetii* prevalence and bacterial load in goat excreta, Emerging Infectious Diseases. (Submitted).

[65] Hohle M., Jorgensen E., O'Neill P.D., Inference in disease transmission experiments by using stochastic epidemic models, Journal of the Royal Statistical Society Series C-Applied Statistics. (2005) 54:349-366.

[66] Huebner R.J., Bell J.A., Q fever studies in Southern California; summary of current results and a discussion of possible control measures, J Am Med Assoc. (1951) 145:301-305; passim.

[67] Hunink J.E., Veenstra T., van der Hoek W., Droogers P., Q fever transmission to humans and local environmental conditions. RIVM, 2010. Available online: http://www.rivm.nl/cib/binaries/Q-fever%20transmission_tcm92-66769.pdf

[68] Jager C., Willems H., Thiele D., Baljer G., Molecular characterization of *Coxiella burnetii* isolates, Epidemiol Infect. (1998) 120:157-164.

[69] James A., Pitchford J.W., Plank M.J., An event-based model of superspreading in epidemics, Proceedings of the Royal Society B: Biological Sciences. (2007) 274:741-747.

[70] John T.J., Samuel R., Herd immunity and herd effect: new insights and definitions, Eur J Epidemiol. (2000) 16:601-606.

[71] Kao R.R., Roberts M.G., Ryan T.J., A model of bovine tuberculosis control in domesticated cattle herds, Proc Biol Sci. (1997) 264:1069-1076.

[72] Karagiannis I., Schimmer B., Van Lier A., Timen A., Schneeberger P., Van Rotterdam B., De Bruin A., Wijkmans C., Rietveld A., Van Duynhoven Y., Investigation of a Q fever outbreak in a rural area of The Netherlands, Epidemiol Infect. (2009) 137:1283-1294.

[73] Kazar J., *Coxiella burnetii* infection, Ann N Y Acad Sci. (2005) 1063:105-114.

[74] Keeling M.J., Woolhouse M.E., Shaw D.J., Matthews L., Chase-Topping M., Haydon D.T., Cornell S.J., Kappey J., Wilesmith J., Grenfell B.T., Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape, Science. (2001) 294:813-817.

[75] Keeling M.J., Woolhouse M.E., May R.M., Davies G., Grenfell B.T., Modelling vaccination strategies against foot-and-mouth disease, Nature. (2003) 421:136-142.

[76] Keeling M.J., Models of foot-and-mouth disease, Proc Biol Sci. (2005) 272:1195-1202.

[77] Keeling M.J., Rohani P., Modeling infectious diseases in Humans and animals, Princeton University Press, Princeton, 2008.

[78] Kersh G.J., Wolfe T.M., Fitzpatrick K.A., Candee A.J., Oliver L.D., Patterson N.E., Self J.S., Priestley R.A., Loftis A.D., Massung R.F., Presence of *Coxiella burnetii* DNA in the environment of the United States, 2006 to 2008, Appl Environ Microbiol 76:4469-4475.

[79] Kittelberger R., Mars J., Wibberley G., Sting R., Henning K., Horner G.W., Garnett K.M., Hannah M.J., Jenner J.A., Piggott C.J., O'Keefe J.S., Comparison of the Q-fever complement fixation test and two commercial enzyme-linked immunosorbent assays for the detection of serum antibodies against *Coxiella* burnetti (Q-fever) in ruminants : recommendations for use of serological tests on imported animals in New Zealand, N Z Vet J. (2009) 57:262-268.

[80] Kobilinsky A., Les plans factoriels, In: Droesbeke, J.-J., Fine, J., Saporta, G. Plans d'expériences : applications à l'entreprise, Editions Technip, Paris, 1997, pp. 69–209.

[81] Kruszewska D., Tylewska-Wierzbanowska S., Isolation of *Coxiella burnetii* from bull semen, Res Vet Sci. (1997) 62:299-300.

[82] Lamboni M., Makowski D., Lehuger S., Gabrielle B., Monod H., Multivariate global sensitivity analysis for dynamic crop models, Field Crops Research. (2009) 113:312-320.

[83] Lamboni M., Monod H., Multisensi: Multivariate Sensitivity Analysis, 2010.

[84] Lamboni M., Monod H., Makowski D., Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models, Reliability and Engineering System Safety (Submitted).

[85] Lanzas C., Warnick L.D., James K.L., Wright E.M., Wiedmann M., Grohn Y.T., Transmission dynamics of a multidrug-resistant Salmonella typhimurium outbreak in a dairy farm, Foodborne Pathog Dis 7:467-474.

[86] Lanzas C., Brien S., Ivanek R., Lo Y., Chapagain P.P., Ray K.A., Ayscue P., Warnick L.D., Grohn Y.T., The effect of heterogeneous infectious period and contagiousness on the dynamics of Salmonella transmission in dairy cattle, Epidemiol Infect. (2008) 136:1496-1510.

[87] Le Menach A., Vergu E., Grais R.F., Smith D.L., Flahault A., Key strategies for reducing spread of avian influenza among commercial poultry holdings: lessons for transmission to humans, Proc Biol Sci. (2006) 273:2467-2475.

[88] Lekone P.E., Finkenstadt B.F., Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study, Biometrics. (2006) 62:1170-1177.

[89] Lipsitch M., Cohen T., Cooper B., Robins J.M., Ma S., James L., Gopalakrishna G., Chew S.K., Tan C.C., Samore M.H., Fisman D., Murray M., Transmission dynamics and control of severe acute respiratory syndrome, Science. (2003) 300:1966-1970.

[90] Liu W.C., Matthews L., Chase-Topping M., Savill N.J., Shaw D.J., Woolhouse M.E., Metapopulation dynamics of *Escherichia coli O157* in cattle: an exploratory model, J R Soc Interface. (2007) 4:917-924.

[91] Lloyd-Smith J.O., Schreiber S.J., Kopp P.E., Getz W.M., Superspreading and the effect of individual variation on disease emergence, Nature. (2005) 438:355-359.

[92] Lu Z., Grohn Y.T., Smith R.L., Wolfgang D.R., Van Kessel J.A., Schukken Y.H., Assessing the potential impact of Salmonella vaccines in an endemically infected dairy herd, J Theor Biol. (2009) 259:770-784.

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

[94] Lurette A., Belloc C., Touzeau S., Hoch T., Ezanno P., Seegers H., Fourichon C., Modelling *Salmonella* spread within a farrow-to-finish pig herd, Vet Res. (2008) 39:49.

[95] Lurette A., Touzeau S., Lamboni M., Monod H., Sensitivity analysis to identify key parameters influencing *Salmonella* infection dynamics in a pig batch, J Theor Biol. (2009) 258:43-52.

[96] Lyytikainen O., Ziese T., Schwartlander B., Matzdorff P., Kuhnhen C., Jager C., Petersen L., An outbreak of sheep-associated Q fever in a rural community in Germany, Eur J Epidemiol. (1998) 14:193-199.

[97] M. J. Keeling, Rohani P., Modeling ingectious diseases in Humans and animals, Princeton University Press, Princeton, 2008.

[98] M. Plommet M.C., J. Gestin, G. Renoux Fièvre Q expérimentale des bovins, Ann. Rech. vétér. (1973) 4:325-346.

[99] Makowski D., Wallach D., Bayesian methods for parameter estimation and data assimilation with crop models [on line] (2006).
http://www.modelia.org/moodle/file.php/3/minicourse_bayesian/BayesianMethodsPart1.ppt

[100] Maltezou H.C., Raoult D., Q fever in children, Lancet Infect Dis. (2002) 2:686-691.

[101] Marce C., Ezanno P., Weber M.F., Seegers H., Pfeiffer D.U., Fourichon C., Invited review: modeling within-herd transmission of M*ycobacterium avium subspecies paratuberculosis* in dairy cattle: a review, J Dairy Sci 93:4455-4470.

[102] Mariner J.C., McDermott J., Heesterbeek J.A., Thomson G., Martin S.W., A model of contagious bovine pleuropneumonia transmission dynamics in East Africa, Prev Vet Med. (2006) 73:55-74.

[103] Marrie T.J., Q fever - a review, Can. Vet. J. (1990) 31:555-563.

[104] Matthews L., Woolhouse M., New approaches to quantifying the spread of infection, Nat Rev Microbiol. (2005) 3:529-536.

[105] Matthews L., Low J.C., Gally D.L., Pearce M.C., Mellor D.J., Heesterbeek J.A., Chase-Topping M., Naylor S.W., Shaw D.J., Reid S.W., Gunn G.J., Woolhouse M.E., Heterogeneous shedding of *Escherichia coli O157* in cattle and its implications for control, Proc Natl Acad Sci USA. (2006) 103:547-552.

[106] Matthews L., McKendrick I.J., Ternent H., Gunn G.J., Synge B., Woolhouse M.E., Super-shedding cattle and the transmission dynamics of *Escherichia coli O157*, Epidemiol Infect. (2006) 134:131-142.

[107] Maurin M., Raoult D., Q fever, Clin Microbiol Rev. (1999) 12:518-553.

[108] McCaughey C., McKenna J., McKenna C., Coyle P.V., O'Neill H.J., Wyatt D.E., Smyth B., Murray L.J., Human seroprevalence to *Coxiella burnetii* (Q fever) in Northern Ireland, Zoonoses Public Health. (2008) 55:189-194.

[109] McQuiston J.H., Childs J.E., Q fever in humans and animals in the United States, Vector Borne Zoonotic Dis. (2002) 2:179-191.

[110] McQuiston J.H., Childs J.E., Thompson H.A., Q fever, J Am Vet Med Assoc. (2002) 221:796-799.

[111] Mediannikov O., Fenollar F., Socolovschi C., Diatta G., Bassene H., Molez J.F., Sokhna C., Trape J.F., Raoult D., *Coxiella burnetii* in humans and ticks in rural Senegal, PLoS Negl Trop Dis. (2010) 4:e654.

[112] Milazzo A., Hall R., Storm P.A., Harris R.J., Winslow W., Marmion B.P., Sexually transmitted Q fever, Clin Infect Dis. (2001) 33:399-402.

[113] Monod H., Makowski D., Naud C., Uncertainty and sensitivity analysis for crop models, In: D. Wallach, D. Makowski, J.W. Jones (Eds), Working with Dynamic Crop Models, Evaluation, Analysis, Parameterization, and Applications, Elsevier, Amsterdam, 2006, pp. 55-100.

[114] Morton A., Finkenstadt B.F., Discrete time modelling of disease incidence time series by using Markov chain Monte Carlo methods, Journal of the Royal Statistical Society Series C-Applied Statistics. (2005) 54:575-594.

[115] National Agricultural Biosecurity Center, Kansas State University [on line] http://nabc.ksu.edu/content/factsheets/category/Q%20Fever [consulted September 2010].

[116] Nielsen S.S., Toft N., Jorgensen E., Bibby B.M., Bayesian mixture models for within-herd prevalence estimates of bovine paratuberculosis based on a continuous ELISA response, Prev Vet Med. (2007) 81:290-305.

[117] Nokes D.J., Swinton J., The control of childhood viral infections by pulse vaccination, IMA J Math Appl Med Biol. (1995) 12:29-53.

[118] Norlander L., Q fever epidemiology and pathogenesis, Microbes and Infection. (2000) 2:417-424.

[119] O'Neill P.D., Roberts G.O., Bayesian inference for partially observed stochastic epidemics, Journal of the Royal Statistical Society Series a-Statistics in Society. (1999) 162:121-129.

[120] Palmer N.C., Kierstead M., Key D.W., Williams J.C., Peacock M.G., Vellend H., Placentitis and Abortion in Goats and Sheep in Ontario Caused by *Coxiella burnetii*, Can Vet J. (1983) 24:60-61.

[121] Panaiotov S., Ciccozzi M., Brankova N., Levterova V., Mitova-Tiholova M., Amicosante M., Rezza G., Kantardjiev T., An outbreak of Q fever in Bulgaria, Ann Ist Super Sanita. (2009) 45:83-86.

[122] Panning M., Kilwinski J., Greiner-Fischer S., Peters M., Kramme S., Frangoulidis D., Meyer H., Henning K., Drosten C., High throughput detection of *Coxiella burnetii* by real-time PCR with internal control system and automated DNA preparation, BMC Microbiol. (2008) 8:77.

[123] Pasanisi A., Aide à la décision dans la gestion des parcs de compteurs d'eau potable Ph.D. thesis, ENGREF, Paris, 2004.

[124] Pinsky R.L., Fishbein D.B., Greene C.R., Gensheimer K.F., An outbreak of cat-associated Q fever in the United States, J Infect Dis. (1991) 164:202-204.

[125] Plommet M., Capponi M., Gestin J., Renoux G., Fièvre Q expérimentale des bovins, Ann Rech Vet. (1973) 4:325-346.

[126] Plummer M., Best N., Cowles K., Vines K., Coda: output analysis and diagnostics for MCMC, R package version 0.13-4, 2009.

[127] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2009.

[128] Raoult D., Tissot-Dupont H., Foucault C., Gouvernet J., Fournier P.E., Bernit E., Stein A., Nesri M., Harle J.R., Weiller P.J., Q fever 1985-1998. Clinical and epidemiologic features of 1,383 infections, Medicine (Baltimore). (2000) 79:109-123.

[129] Read A.J., Erickson S., Harmsen A.G., Role of CD4+ and CD8+ T cells in clearance of primary pulmonary infection with *Coxiella burnetii*, Infect Immun. (2010) 78:3019-3026.

[130] Rivot E., Prevost E., Parent E., Bagliniere J.L., A Bayesian state-space modelling framework for fitting a salmon stage-structured population dynamic model to multiple time series of field data, Ecological Modelling. (2004) 179:463-485.

[131] Rodolakis A., Berri M., Hechard C., Caudron C., Souriau A., Bodier C.C., Blanchard B., Camuset P., Devillechaise P., Natorp J.C., Vadet J.P., Arricau-Bouvery N., Comparison of *Coxiella burnetii* shedding in milk of dairy bovine, caprine, and ovine herds, J Dairy Sci. (2007) 90:5352-5360.

[132] Rodolakis A., Q Fever in dairy animals, Ann N Y Acad Sci. (2009) 1166:90-93.

[133] Rodolakis A., Clement P., Cochonneau D., Beaudeau F., Sarradin P., Guatteo R., Investigation of humoral and cellular immunity of dairy cattle after one or two year of vaccination with a phase I *Coxiella* vaccine, Procedia in Vaccinology. (2009) 1:85-88.

[134] Rodriguez N.F., Carranza C., Bolanos M., Perez-Arellano J.L., Gutierrez C., Seroprevalence of *Coxiella burnetii* in domestic ruminants in Gran Canaria Island, Spain, Transbound Emerg Dis. (2010) 57:66-67.

[135] Rousset E., Berri M., Durand B., Dufour P., Prigent M., Delcroix T., Touratier A., Rodolakis A., *Coxiella burnetii* shedding routes and antibody response after outbreaks of Q fever-induced abortion in dairy goat herds, Appl Environ Microbiol. (2009) 75:428-433.

[136] Rousset E., Durand B., Champion J.L., Prigent M., Dufour P., Forfait C., Marois M., Gasnier T., Duquesne V., Thiery R., Aubert M.F., Efficiency of a phase 1 vaccine for the reduction of vaginal *Coxiella burnetii* shedding in a clinically affected goat herd, Clin Microbiol Infect. (2009) 15 Suppl 2:188-189.

[137] Roy C., Rhinotrachéïte Infectieuse Bovine (IBR) [on line] (2007). http://www.agrireseau.qc.ca/bovinslaitiers/documents/Rhinotrach%C3%A9%C3%AFte%20infectieuse%20-%20S%C3%A9minaire%20%28C.%20Roy%29.pdf [consulted 18/10/10].

[138] Ruiz-Fons F., Astobiza I., Barandika J.F., Hurtado A., Atxaerandio R., Juste R.A., Garcia-Perez A.L., Seroepidemiological study of Q fever in domestic ruminants in semi-extensive grazing systems, BMC Vet Res 6:3.

[139] Sadecky E., Brezina R., Kazar J., Urvolgyi J., Immunization against Q-fever of naturally infected dairy cows, Acta Virol. (1975) 19:486-488.

[140] Saltelli A., Chan K., Scott E.M., Sensitivity Analysis, Wiley, Chichester, 2000.

[141] Schimmer B., Dijkstra F., Vellema P., Schneeberger P.M., Hackert V., ter Schegget R., Wijkmans C., van Duynhoven Y., van der Hoek W., Sustained intensive transmission of Q fever in the south of the Netherlands, 2009, Euro Surveill. (2009) 14.

[142] Schimmer B., Ter Schegget R., Wegdam M., Zuchner L., de Bruin A., Schneeberger P.M., Veenstra T., Vellema P., van der Hoek W., The use of a geographic information system to identify a dairy goat farm as the most likely source of an urban Q-fever outbreak, BMC Infect Dis. (2010) 10:69.

[143] Sharomi O., Podder C.N., Gumel A.B., Mahmud S.M., Rubinstein E., Modelling the Transmission Dynamics and Control of the Novel 2009 Swine Influenza (H1N1) Pandemic, Bull Math Biol. (2010) in press.

[144] Smith R.L., Sanderson M.W., Renter D.G., Larson R., White B., A stochastic risk-analysis model for the spread of bovine viral diarrhea virus after introduction to naive cow-calf herds, Prev Vet Med 95:86-98.

[145] Streftaris G., Gibson G.J., Bayesian analysis of experimental epidemics of foot-and-mouth disease, Proc Biol Sci. (2004) 271:1111-1117.

[146] Suppo C., Naulin J.M., Langlais M., Artois M., A modelling approach to vaccination and contraception programmes for rabies control in fox populations, Proc Biol Sci. (2000) 267:1575-1582.

[147] Svraka S., Toman R., Skultety L., Slaba K., Homan W.L., Establishment of a genotyping scheme for *Coxiella burnetii*, FEMS Microbiol Lett. (2006) 254:268-274.

[148] Szmaragd C., Wilson A.J., Carpenter S., Wood J.L., Mellor P.S., Gubbins S., The spread of bluetongue virus serotype 8 in Great Britain and its control by vaccination, PLoS One 5:e9353.

[149] Taurel A.F., Guatteo R., Joly A., Seegers H., Beaudeau F., Q fever within-herd seroprevalence in infected dairy herds: assessment using an ELISA applied to bulk tank milk [on line] (2010). http://www.svepm.org.uk/posters/2010/Taurel_Q%20fever%20within-herd%20seroprevalence%20in%20infected%20dairy%20herd_assessment%20using%20an%20ELISA%20applied%20to%20bulk%20tan.pdf [consulted 17/08/2010].

[150] Tigertt W.D., Benenson A.S., Gochenour W.S., Airborne Q fever, Bacteriol Rev. (1961) 25:285-293.

[151] Tissot-Dupont H., Amadei M.A., Nezri M., Raoult D., Wind in November, Q fever in December, Emerg Infect Dis. (2004) 10:1264-1269.

[152] Tissot Dupont H., Raoult D., Brouqui P., Janbon F., Peyramond D., Weiller P.J., Chicheportiche C., Nezri M., Poirier R., Epidemiologic features and clinical presentation of acute Q fever in hospitalized patients: 323 French cases, Am J Med. (1992) 93:427-434.

[153] To H., Htwe K.K., Kako N., Kim H.J., Yamaguchi T., Fukushi H., Hirai K., Prevalence of *Coxiella burnetii* Infection in Dairy Cattle with Reproductive Disorders, The Journal of Veterinary Medical Science. (1998) 60:859-861.

[154] Toledo A., Jado I., Olmeda A.S., Casado-Nistal M.A., Gil H., Escudero R., Anda P., Detection of *Coxiella burnetii* in ticks collected from Central Spain, Vector Borne Zoonotic Dis. (2009) 9:465-468.

[155] Treanor J.J., Johnson J.S., Wallen R.L., Cilles S., Crowley P.H., Cox J.J., Maehr D.S., White P.J., Plumb G.E., Vaccination strategies for managing brucellosis in Yellowstone bison, Vaccine 28 Suppl 5:F64-72.

[156] Turner J., Bowers R.G., Begon M., Robinson S.E., French N.P., A semi-stochastic model of the transmission of Escherichia coli O157 in a typical UK dairy herd: dynamics, sensitivity analysis and intervention/prevention strategies, J Theor Biol. (2006) 241:806-822.

[157] Turner J., Bowers R.G., Begon M., Robinson S.E., French N.P., A semi-stochastic model of the transmission of Escherichia coli O157 in a typical UK dairy herd: Dynamics, sensitivity analysis and intervention/prevention strategies, Journal of Theoretical Biology. (2006) 241:806-822.

[158] Valleron A.J., Les rôles de la modélisation en épidémiologie, C.R. Acad. Sci. Paris. (2000) 323:429-433.

[159] van der Hoek W., Dijkstra F., Schimmer B., Schneeberger P.M., Vellema P., Wijkmans C., ter Schegget R., Hackert V., van Duynhoven Y., Q fever in the Netherlands: an update on the epidemiology and control measures, Euro Surveill. (2010) 15.

[160] Van Effelterre T., Moore M.R., Fierens F., Whitney C.G., White L., Pelton S.I., Hausdorff W.P., A dynamic model of pneumococcal infection in the United States: implications for prevention through vaccination, Vaccine 28:3650-3660.

[161] Vellema P., Van den Brom R., Dercksen D., Moll L., Roest H.J., Q fever in the Netherlands: One Health in relation to Q fever, in humans and animals [on line] (2010). www.minlnv.nl/txmpub/files/?p_file_id=2000277 [consulted 17/08/2010].

[162] Viet A.F., Fourichon C., Seegers H., Review and critical discussion of assumptions and modelling options to study the spread of the bovine viral diarrhoea virus (BVDV) within a cattle herd, Epidemiol Infect. (2007) 135:706-721.

[163] Vynnycky E., White R.G., An introduction to infectious disease modelling, Oxford University Press, New York, 2010.

[164] Wallensten A., Moore P., Webster H., Johnson C., van der Burgt G., Pritchard G., Ellis-Iversen J., Oliver I., Q fever outbreak in Cheltenham, United Kingdom, in 2007 and the use of dispersion modelling to investigate the possibility of airborne spread, Euro Surveill. (2010) 15.

[165] Weber M.F., Nielen M., Velthuis A.G., van Roermund H.J., Milk quality assurance for paratuberculosis: simulation of within-herd infection dynamics and economics, Vet Res. (2008) 39:12.

[166] Webster J.P., Lloyd G., Macdonald D.W., Q fever (*Coxiella burnetii*) reservoir in wild brown rat (Rattus norvegicus) populations in the UK, Parasitology. (1995) 110 (Pt 1):31-35.

[167] Welsh H.H., Lennette E.H., Abinanti F.R., Winn J.F., Air-borne transmission of Q fever : the role of parturition in the generation of infective aerosols, Ann NY Acad Sci. (1958) 70:528-540.

[168] Woldehiwet Z., Q fever (coxiellosis): epidemiology and pathogenesis, Res Vet Sci. (2004) 77:93-100.

[169] Woolhouse M.E., Dye C., Etard J.F., Smith T., Charlwood J.D., Garnett G.P., Hagan P., Hii J.L., Ndhlovu P.D., Quinnell R.J., Watts C.H., Chandiwana S.K., Anderson R.M., Heterogeneities in the

transmission of infectious agents: implications for the design of control programs, Proc Natl Acad Sci U S A. (1997) 94:338-342.

[170] World Health Organization [on line] http://www.who.int/zoonoses/control_neglected_zoonoses/en/index.html [consulted September 2010].

[171] Zhang G., Russell-Lodrigue K.E., Andoh M., Zhang Y., Hendrix L.R., Samuel J.E., Mechanisms of vaccine-induced protective immunity against *Coxiella burnetii* infection in BALB/c mice, J Immunol. (2007) 179:8372-8380.

[172] Zhang X.S., Chase-Topping M.E., McKendrick I.J., Savill N.J., Woolhouse M.E., Spread of *E. coli O157* infection among Scottish cattle farms: Stochastic models and model selection, Epidemics 2:11-20.

Summary in French

Resume en français

***Chapitre 1* : Introduction**

Coxiella burnetii est une bactérie intracellulaire stricte responsable d'une zoonose mondialement répandue, la fièvre Q. Le contrôle de cette infection est un enjeu crucial :

- il s'agit tout d'abord d'un problème de santé publique : bien qu'asymptomatique dans plus de 60% des cas, la fièvre Q peut entraîner chez l'Homme des signes cliniques graves tels que pneumonies, hépatites, endocardites ou avortements. Aux Pays-Bas, une épidémie massive sévit depuis 2007 : plus de 3500 cas humains y ont déjà été confirmés. Les ruminants sont reconnus comme la principale source de contamination de l'Homme : les animaux infectés excrètent de grandes quantités de bactéries dans l'environnement via les fèces, l'urine, le lait et surtout les produits de parturition (placenta notamment). *C. burnetii* s'avérant très résistante dans l'environnement, on la retrouve soit sous forme d'aérosols soit dans la poussière environnante. Ces deux supports sont les principales sources d'infection pour l'Homme.
- la fièvre Q constitue également un problème de santé animale : chez les ruminants, cette infection peut entraîner des troubles de la reproduction tels qu'avortements, métrites ou infertilité, engendrant des pertes économiques importantes pour les élevages atteints.

Il apparaît donc essentiel de lutter contre la propagation de *C. burnetii* dans les troupeaux de ruminants domestiques afin d'améliorer les performances de ces élevages et de limiter le risque zoonotique. Récemment, l'Agence Européenne de Sécurité des Aliments (EFSA) a souligné le besoin, d'une part de quantifier les paramètres clés de l'infection, notamment les taux de transmission (i) au sein des troupeaux de ruminants, (ii) entre ces troupeaux et (iii) des populations animales à l'Homme, et d'autre part d'évaluer l'efficacité des stratégies de contrôle. Ce travail de thèse s'inscrit dans ce contexte général. Il a pour objectif d'améliorer la compréhension de la propagation de *C. burnetii* au sein d'un troupeau bovin laitier afin de proposer des stratégies de contrôle efficaces. Pour cela, nous avons développé une approche par modélisation: pour des raisons éthiques, logistiques et financières, des études observationnelles en troupeaux infectés ne peuvent pas être mises en œuvre sur le long terme afin d'étudier toutes les stratégies possibles de contrôle de l'infection dans l'ensemble des situations épidémiologiques possibles. Il nous a donc paru pertinent de développer un modèle mathématique de la propagation de la bactérie permettant de suivre chaque paramètre de l'infection et de comparer l'efficacité de différentes stratégies de maîtrise ex ante.

Trois étapes de travail nous ont permis d'atteindre notre objectif. Un modèle représentant la propagation de *C. burnetii* au sein d'un troupeau bovin laitier a tout d'abord été conceptualisé et

ses principaux paramètres épidémiologiques estimés à partir de données de terrain, en utilisant une approche Bayésienne (chapitre 3). Il s'avère que ces données suggéraient l'existence d'une forte hétérogénéité d'excrétion au sein des troupeaux infectés (les voies et durées d'excrétion de même que les concentrations en bactéries excrétées sont très variables d'une vache à l'autre et d'un moment à l'autre pour une même vache) et que la présence d'hétérogénéité au sein d'une population influence très souvent les dynamiques d'infection. Nous avons donc décidé de représenter explicitement dans une version plus complète du modèle les voies d'excrétion et les quantités de bactéries excrétées. Nous avons ensuite réalisé une analyse de sensibilité afin d'identifier les paramètres dont la variation influençait le plus la dynamique d'infection (chapitre 4). Enfin, nous avons représenté dans le modèle différentes stratégies de vaccination et comparé leurs efficacités respectives par simulation (chapitre 5).

<u>Chapitre 2</u> : Elaboration d'un modèle de propagation de C. burnetii au sein d'un troupeau bovin laitier et estimation de ses paramètres principaux à partir de données de terrain

La 1ère partie de ce chapitre décrit les principales étapes à suivre pour construire un modèle épidémiologique et le confronter à des données de terrain. Nous nous focalisons dans une 2ème partie sur le modèle développé pour représenter la propagation de *C. burnetii* au sein d'un troupeau bovin. Comme les données de terrain sont essentielles à sa conceptualisation et à l'inférence de ses paramètres, nous avons tout d'abord décrit le jeu de données que nous avons utilisé. 235 vaches de cinq troupeaux naturellement infectés et ne présentant pas de signes cliniques attribuables à *C. burnetii* ont été prélevées entre une et cinq fois sur une période d'un mois. A chaque temps de prélèvement, un test sérologique ainsi que 3 tests PCR en temps réel (un sur lait, un sur mucus vaginal et un sur fèces) ont été réalisés. Il nous a donc été possible de définir le statut vis-à-vis de l'infection de chaque vache à chaque pas de temps. Au total, nous disposions pour 145 vaches de leur statut hebdomadaire et pour 89 autres, d'un à quatre statuts au cours du mois d'étude. Nous présentons ensuite le modèle élaboré : il s'agit d'un modèle SIR (Sensibles-Infectieux-Retirés de la chaîne de transmission) modifié, caractérisé par 2 classes de I ($I-$ et $I+$, pour excréteurs séronégatifs et séropositifs respectivement) et par des transitions dans les deux sens entre S et $I-$ et entre $I+$ et R. Etant donné que la contamination d'un animal se fait par inhalation, la probabilité d'infection (transition de S à $I-$) est supposée dépendre de la charge bactérienne dans l'environnement. Ce modèle est stochastique, individu-centré et en temps discret avec un pas de temps d'une semaine. Enfin, nous exposons dans la 3ème partie de ce chapitre, l'estimation des paramètres épidémiologiques de ce modèle à partir du jeu de données précédemment évoqué : une approche Bayésienne a été

privilégiée car elle permettait de prendre en compte les connaissances disponibles a priori et de gérer les données manquantes, ainsi que l'incertitude des observations due à l'imperfection des tests diagnostics. Par l'utilisation d'algorithmes d'estimation par Chaînes de Markov de Monte-Carlo, nous avons obtenu les distributions *a posteriori* des probabilités de transitions entre états de santé et de la charge bactérienne environnementale Les résultats ont montré que certains troupeaux étaient caractérisés par un faible risque d'infection alors que pour d'autres, ce risque, tout comme la probabilité d'excrétion intermittente étaient modérés. De plus, les excréteurs séronégatifs (*I-*) excrétaient moins longtemps que les excréteurs séropositifs (*I+*).

<u>Chapitre 3</u> : Représentation de l'hétérogénéité d'excrétion dans le modèle de propagation de C. burnetii et identification des paramètres influençant le plus la dynamique d'infection

La 1ère partie de chapitre définit la notion d'hétérogénéité en population d'hôtes : elle présente l'impact de cette hétérogénéité sur la dynamique d'infection de nombreux pathogènes, ses implications en termes de contrôle et la manière dont elle est prise en compte dans la modélisation épidémiologique. La 2ème partie se focalise sur l'hétérogénéité d'excrétion de *C. burnetii* en troupeaux bovins : la variabilité des voies d'excrétion et des concentrations de bactéries excrétées observée dans le jeu de données, détaillé au chapitre 2 y est décrite. La 3ème partie du chapitre, rappelle les différentes méthodes d'analyse de sensibilité existantes dans la littérature. L'analyse de sensibilité est en effet une étape cruciale du processus de modélisation car elle permet l'identification des paramètres influençant majoritairement la dynamique d'infection. Les paramètres identifiés revêtent une double importance : ils doivent être très précisément estimés si l'on veut améliorer les capacités prédictives du modèle et ils sont les cibles à privilégier pour la définition de mesures de contrôle efficaces. Enfin, la 4ème partie expose le modèle représentant la propagation intra troupeau de *C. burnetii* en incluant l'hétérogénéité d'excrétion, et décrit l'analyse de sensibilité que nous avons employée. Pour comparer l'influence des paramètres épidémiologiques sur différentes sorties temporelles du modèle, nous avons en effet effectué une Analyse en Composantes Principales (ACP) suivie d'une ANOVA. Nous avons ainsi montré que les paramètres les plus influents étaient les distributions de probabilité gouvernant les quantités de bactéries excrétées, principalement dans le mucus vaginal et les fèces, les caractéristiques de *C. burnetii* dans l'environnement (i.e. sa survie et la fraction de bactéries excrétées atteignant l'environnement) ainsi que les caractéristiques physiologiques liées à l'intermittence de l'excrétion et à la transition d'un type d'excréteur (*I-*) à l'autre (*I+*).

***Chapitre 4* : Comparaison de l'efficacité de trois stratégies de vaccination en troupeaux infectés**

Dans la 1ère partie de ce chapitre, nous avons étudié quelques exemples de stratégies vaccinales et la manière dont elles sont représentées dans les modèles épidémiologiques de la littérature. Dans la 2nde partie, nous nous sommes intéressés à l'évaluation de l'efficacité relative de différentes stratégies de vaccination contre *C. burnetii* en troupeaux bovins infectés. Trois indicateurs temporels de la dynamique d'infection (i.e. prévalence en excréteurs, charge bactérienne environnementale et nombre d'avortements) ainsi que la probabilité d'extinction de l'infection ont en effet été simulés pour trois scénarii de vaccination ainsi qu'un scénario témoin sans stratégie de contrôle. Pour tous les scénarii avec vaccination, les valeurs de ces trois indicateurs ont baissé durant les premières années du programme de vaccination. Cependant, une vaccination d'une durée limitée (trois ans seulement) était souvent insuffisante pour éradiquer l'infection : l'arrêt du programme de vaccination entraînait donc une reprise de la propagation de l'infection. De plus, à la mise en place du programme, la vaccination des vaches et des génisses était préférable à celle des génisses seulement. Dans ce dernier cas, les indicateurs de la dynamique d'infection décroissaient plus lentement et le taux d'extinction de l'infection était deux fois plus faible que lorsque vaches et génisses étaient vaccinées.

***Chapitre 5* : Discussion générale**

La 1ère partie de ce chapitre reprend les résultats majeurs de cette thèse : nous avons mis en évidence une variabilité de la dynamique d'infection entre troupeaux *a priori* similaires (i.e. troupeaux infectés par *C. burnetii* sans signes cliniques attribuables à la maladie) et une variabilité de l'excrétion entre animaux. Nous avons de plus identifié les paramètres influençant le plus la dynamique d'infection et évalué l'efficacité relative de trois stratégies de vaccination. Dans une 2ème partie, nous présentons les avantages et limites de l'approche de modélisation. La 3ème partie discute des données disponibles et de celles qui seraient nécessaires pour préciser la conceptualisation et la paramétrisation du modèle ainsi que sa validation : la réponse immunitaire cellulaire des animaux, les premières étapes de la propagation de *C. burnetii* au sein d'un troupeau, ainsi que la viabilité et la quantification de la bactérie dans l'environnement mériteraient en effet d'être étudiées. Enfin, les implications et les perspectives de cette thèse sont énoncées dans la 4ème partie de ce chapitre. En conclusion, outre la première quantification des paramètres épidémiologiques de la propagation de *C. burnetii* dans un troupeau et de nouvelles interprétations des mécanismes impliqués, ce travail fournit une aide à la priorisation des besoins de recherche et à la définition de mesures efficaces pour contrôler la fièvre Q en troupeaux bovins laitiers.

Abstract

Q fever is a worldwide zoonosis caused by *Coxiella burnetii* which induces reproductive disorders in livestock. Ruminants are also recognized as the most important source of human infection. Therefore, the control of this infection in cattle is crucial to limit both the infection in livestock and the zoonotic risk. The objective of this thesis was to better understand the natural course of the infection within dairy cattle herds in order to propose effective control measures. A stochastic individual-based model in discrete time was conceptualised to represent the *C. burnetii* spread within a dairy herd. Its main epidemiological parameters were assessed from field data using a Bayesian approach. As a great heterogeneity between shedder cows, known to impact infection dynamics, has been described, the shedding routes and levels were explicitly represented in a variant of the first model. The most influential parameters of the infection dynamics, identified through a sensitivity analysis, were the levels of shedding, the characteristics of the bacterium in the environment and some physiological features of cows. Lastly, the long-term effectiveness of three different vaccination strategies in reducing the shedders prevalence, the number of abortions, the environmental bacterial load, and in leading to infection extinction was tested by simulation. A 10-year vaccination programme for both cows and heifers was found to be the most effective one. Besides providing a better understanding of *C. burnetii* infection dynamics, this work can help prioritizing needs of research and designing effective control programmes for Q fever in cattle.

Résumé

La fièvre Q est une zoonose mondialement répandue due à *Coxiella burnetii*. Elle peut engendrer des troubles de la reproduction chez les ruminants. De plus, ces derniers constituent la principale source d'infection pour l'Homme. Il est donc nécessaire de lutter contre la propagation de *C. burnetii* en troupeaux bovins pour améliorer les performances de ces élevages et limiter le risque zoonotique. L'objectif de cette thèse a été de mieux comprendre la propagation de l'infection au sein d'un troupeau bovin laitier, afin de mieux la contrôler. Un modèle épidémiologique stochastique, individu-centré et en temps discret représentant la propagation intra-troupeau de *C. burnetii* a été développé. Ses paramètres ont été estimés à partir de données de terrain en utilisant une approche Bayésienne. Une forte hétérogénéité entre vaches excrétrices ayant été rapportée, les voies et niveaux d'excrétion ont été explicitement représentés dans une variante du premier modèle. Les paramètres influençant le plus la dynamique d'infection, identifiés par une analyse de sensibilité, étaient les niveaux d'excrétion, les caractéristiques de la bactérie dans l'environnement et certains traits physiologiques des animaux. Enfin, trois stratégies de vaccination ont été représentées dans le modèle et leurs efficacités à long terme ont été comparées par simulation. La vaccination des vaches et génisses pendant 10 ans s'est avérée la stratégie la plus efficace. En conclusion, outre une meilleure compréhension de la dynamique d'infection, ce travail fournit une aide à la priorisation des besoins de recherche et à la définition des mesures efficaces pour contrôler la fièvre Q en troupeaux bovins laitiers.